SVM 10 application usage tutorial

Introduction

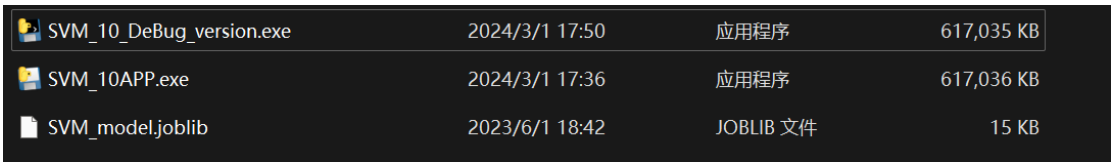
Welcome to the SVM 10 application, a classification model built on the latest research for the molecular subtyping of Hepatocellular Carcinoma (HCC). HCC's high heterogeneity has long been a challenge in treatment, making molecular subtyping essential for personalized anti-tumor therapies.

In our study, we leveraged the critical role of driver genes in cancer development, constructing a classification model with 1192 driver genes from mutational aspects and 233 from driver dysregulation. Through our innovative computational framework, we applied these driver genes to subtype HCC, resulting in two distinct subtypes, CLASS A and CLASS B, with significant differences in survival characteristics.

This tutorial is designed to enable scholars outside the field of computational biology to easily utilize the SVM 10 application, gaining in-depth insights into the molecular subtypes of HCC and their potential therapeutic strategies. We provide detailed steps and instructions to ensure the accurate and straightforward operation of this powerful tool.

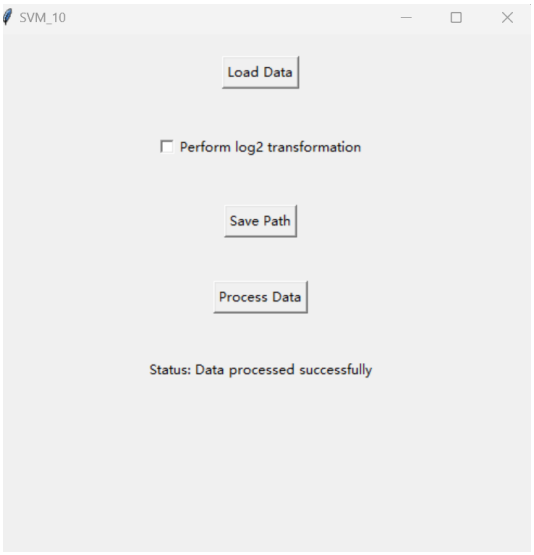
Thank you for using SVM 10, and we look forward to you gaining valuable insights into the molecular subtypes of HCC through this application.

Step 1: Please confirm that the following three files are in the same folder.



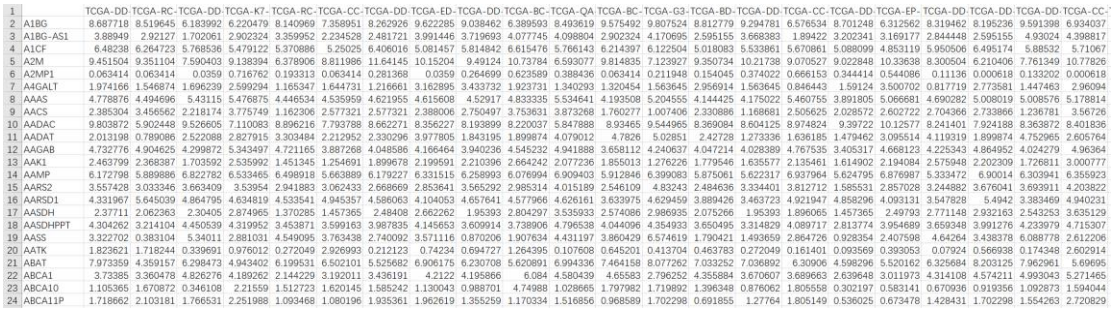
Please note the necessary change in the name of the model file "SVM_model.joblib". Ensure that the model file is in the same folder as the application "SVM_10APP.exe".

Step 2: Open "SVM_10APP.exe" as shown in the figure below (please be patient as the file opening may take some time)



Step 3: Click on "Load Data" and find your data file. Please note that the input file should be a gene expression matrix. Ensure that when inputting the file, it adheres to the following specifications:

1. The input file must be in .csv format.
2. The input data should have genes listed in rows and samples arranged in columns.
3. Make sure your data includes expression data for the following ten genes: ARHGAP11A, CHAF1B, DEPDC1B, ECT2, GINS1, GTSE1, LMNB1, MYBL2, RFC4, TTK.



		TCGA-DD-	TCGA-RC-	TCGA-DD-	TCGA-K7-	TCGA-RC-	TCGA-CC-	TCGA-DD-	TCGA-ED-	TCGA-DD-	TCGA-BC-	TCGA-QA-	TCGA-BC-	TCGA-G3-	TCGA-BD-	TCGA-DD-	TCGA-CC-	TCGA-DD-	TCGA-EP-	TCGA-DD-	TCGA-DD-	TCGA-DD-	TCGA-CC-
2	A1BG	8.687718	8.519645	6.183992	6.220479	8.140969	7.358951	8.262926	9.622285	9.038462	6.389593	8.493619	9.575492	9.807524	8.812779	9.294781	6.576534	8.701248	6.312562	8.319462	8.195236	9.591398	6.934037
3	A1BG-AS1	3.88949	2.92127	1.702061	2.902224	3.359952	2.234528	2.481721	3.991446	3.719693	4.077745	4.098804	2.902324	4.170695	2.595155	3.668383	1.89422	3.202341	3.169177	2.844448	2.595156	4.93024	4.398817
4	A1CF	6.48238	6.264723	5.768536	5.479122	5.370886	5.25025	6.406016	5.081457	5.814842	6.615476	5.766143	6.214397	6.122504	5.018083	5.533861	5.670861	5.088099	4.853119	5.950506	6.495174	5.88532	5.71067
5	A2M	9.451504	9.351104	7.590403	9.138394	6.378906	8.811986	11.64145	10.15204	9.49124	10.73784	6.593077	9.814835	7.123927	9.350734	10.21738	9.070527	9.022848	10.33638	8.300504	6.210406	7.761349	10.77826
6	A2MP1	0.063414	0.063414	0.0359	0.716762	0.193313	0.063414	0.281368	0.0359	0.264699	0.623589	0.388436	0.063414	0.211948	0.154045	0.374022	0.666153	0.344414	0.544086	0.11136	0.000618	0.133202	0.000618
7	A4GALT	1.974166	1.546874	1.696239	2.599294	1.165347	1.644731	1.216661	3.162895	3.433732	1.923731	1.340293	1.320454	1.563645	2.956914	1.563645	0.846443	1.59124	3.500702	0.817719	2.773581	1.447463	2.96094
8	AAAS	4.778876	4.494696	5.43115	5.476875	4.446534	4.535959	4.621955	4.615608	4.52917	4.833335	5.534641	4.193508	5.204555	4.144425	4.175022	5.460755	3.891805	5.066681	4.690282	5.008019	5.008576	5.178814
9	AACS	2.385304	3.456562	2.218174	3.775749	1.162306	2.577321	2.577321	2.388006	2.750497	3.753631	3.873268	1.760277	1.007406	2.330886	1.168681	2.505625	2.028572	2.602722	2.704366	2.733866	1.236781	3.56726
10	AADAC	9.803872	5.902448	9.526605	7.110083	8.896216	7.793788	8.662271	8.356227	8.193899	8.220037	5.847888	8.93465	9.544965	8.369084	8.604125	8.974824	9.39722	10.12577	8.241401	7.924188	8.363872	8.401836
11	AADAT	2.013198	0.789086	2.522088	2.827915	3.303484	2.212952	2.330296	3.977805	1.843195	1.899874	4.079012	4.7826	5.02851	2.42728	1.273336	1.636185	1.479462	3.095514	4.119319	1.899874	4.752965	2.605764
12	AAGAB	4.732776	4.904625	4.299872	5.343497	4.721165	3.887268	4.048586	4.166464	3.940236	4.545232	4.941888	3.658112	4.240637	4.047214	4.028389	4.767535	3.405317	4.668123	4.225343	4.864952	4.024279	4.96364
13	AAK1	2.463799	2.368387	1.703592	2.535992	1.451345	1.254691	1.899678	2.199591	2.210396	2.664242	2.077236	1.855013	1.276226	1.779546	1.635577	2.135461	1.614902	2.194084	2.575948	2.202309	1.726811	3.000777
14	AAMP	6.172798	5.889886	6.822782	6.533465	6.498918	5.663889	6.179227	6.331515	6.258993	6.076994	6.909403	5.912846	6.390083	5.875061	5.622317	6.937964	5.624795	6.876987	5.333472	6.90014	6.303941	6.355923
15	AARS2	3.557428	3.033346	3.663409	3.53954	2.941883	3.062433	2.668669	2.853641	3.565292	2.985314	4.015189	2.546109	4.83243	2.484636	3.334401	3.812712	1.585531	2.857028	3.244882	3.676041	3.693911	4.203822
16	AARSD1	4.331967	5.645039	4.864795	4.634819	4.533541	4.945357	4.586063	4.104053	4.657641	4.577966	4.626161	3.633975	4.629459	3.889426	3.463723	4.921947	4.858296	4.093131	3.547828	5.4942	3.383469	4.940231
17	AASDH	2.37711	2.062363	2.30405	2.874965	1.370285	1.457365	2.48408	2.662262	1.95393	2.804297	3.535933	2.574086	2.986935	2.075266	1.95393	1.896065	1.457365	2.49793	2.771148	2.932163	2.543253	3.635129
18	AASDHPPT	4.304262	3.214104	4.450539	4.319952	3.453871	3.599163	3.987835	4.145653	3.609914	3.738906	4.796538	4.044096	4.354933	3.650495	3.314829	4.089717	2.813774	3.954689	3.659348	3.991276	4.233979	4.715307
19	AASS	3.322702	0.383104	5.34011	2.881031	4.549095	3.763438	2.740092	3.571116	0.870206	1.907634	4.431197	3.860429	6.574619	1.790421	1.493659	2.864726	0.928354	2.407598	4.64264	3.438378	6.088778	2.612206
20	AATK	1.823621	1.718244	0.339691	0.976012	0.272049	2.926993	0.212123	0.74234	0.694727	1.264395	0.107608	0.645201	0.413704	0.463783	0.272049	0.161401	0.093569	0.393053	0.07924	0.566938	0.174348	2.602914
21	ABAT	7.973359	4.359157	6.298473	4.943402	6.199531	6.502101	5.525682	6.906175	6.230708	5.620891	6.994336	7.464158	8.077262	7.033252	7.036892	6.30906	4.598296	5.520162	6.325684	8.203125	7.962961	5.69695
22	ABCA1	3.73385	3.360478	4.826276	4.189262	2.144229	3.192011	3.436191	4.2122	4.195866	6.084	4.580439	4.65583	2.796252	4.355884	3.670607	3.689663	2.639648	3.011973	4.314108	4.574211	4.993043	5.271465
23	ABCA10	1.105365	1.670872	0.346108	2.21559	1.512723	1.620145	1.585242	1.130043	0.988701	4.74988	1.028665	1.797982	1.719892	1.396348	0.876062	1.805558	0.302197	0.583141	0.670936	0.919356	1.092873	1.594044
24	ABCA11P	1.718662	2.103181	1.766531	2.251988	1.093468	1.080196	1.935361	1.962619	1.355259	1.170334	1.516856	0.968589	1.702298	0.691855	1.27764	1.805149	0.536025	0.673478	1.428431	1.702298	1.554263	2.720829

Input data

Step 4: During model construction, we used log2(TPM+1) as the data format. If your data has not undergone log2 transformation, check the "Perform log2 transformation" box. If your data has already undergone log2 transformation, skip this step.

Step 5: Click on "Save Path" to specify the save path for the result file.

Step 6: Click on "Process Data" to execute the data. When the status shows "Data processed successfully" below, the result file will be saved in your specified folder as "subtype_result.csv".

	A	B	C	D
1	Sample	Probability	Probability	Prediction
2	TCGA-DD-	0.72915	0.27085	CLASS A
3	TCGA-RC-	0.031984	0.968016	CLASS B
4	TCGA-DD-	0.135677	0.864323	CLASS B
5	TCGA-K7-	0.020346	0.979654	CLASS B
6	TCGA-RC-	0.618709	0.381291	CLASS A
7	TCGA-CC-	0.1836	0.8164	CLASS B
8	TCGA-DD-	0.208323	0.791677	CLASS B
9	TCGA-ED-	0.903299	0.096701	CLASS A
10	TCGA-DD-	0.479871	0.520129	CLASS B
11	TCGA-BC-	1.77E-05	0.999982	CLASS B
12	TCGA-QA-	2.84E-07	1	CLASS B
13	TCGA-BC-	0.981256	0.018744	CLASS A
14	TCGA-G3-	0.973391	0.026609	CLASS A
15	TCGA-BD-	0.966505	0.033495	CLASS A

Result file

Regarding software malfunction, please open another file named "SVM_10_DeBug_version.exe". This file has two components: program theme and runtime log. Specific error information is provided on the task log page. As shown in the figure below, the error is due to the model file "SVM_model.joblib" being renamed or not being in the same folder as the program file.,