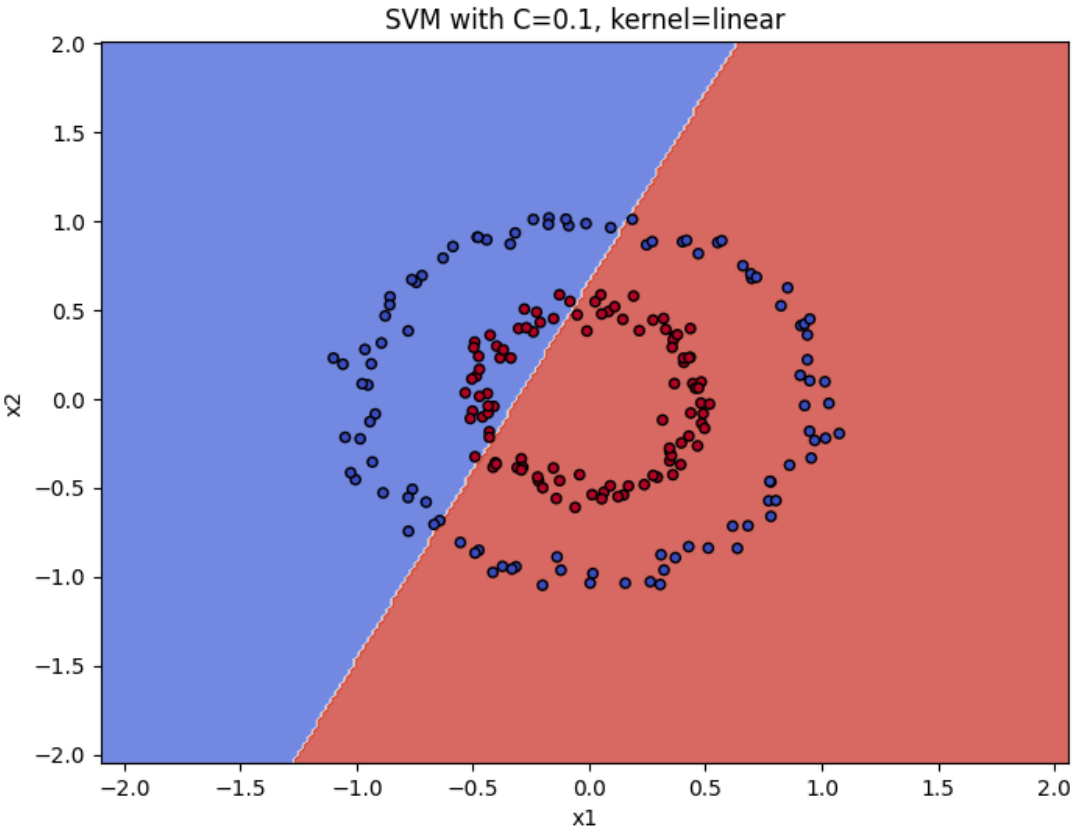
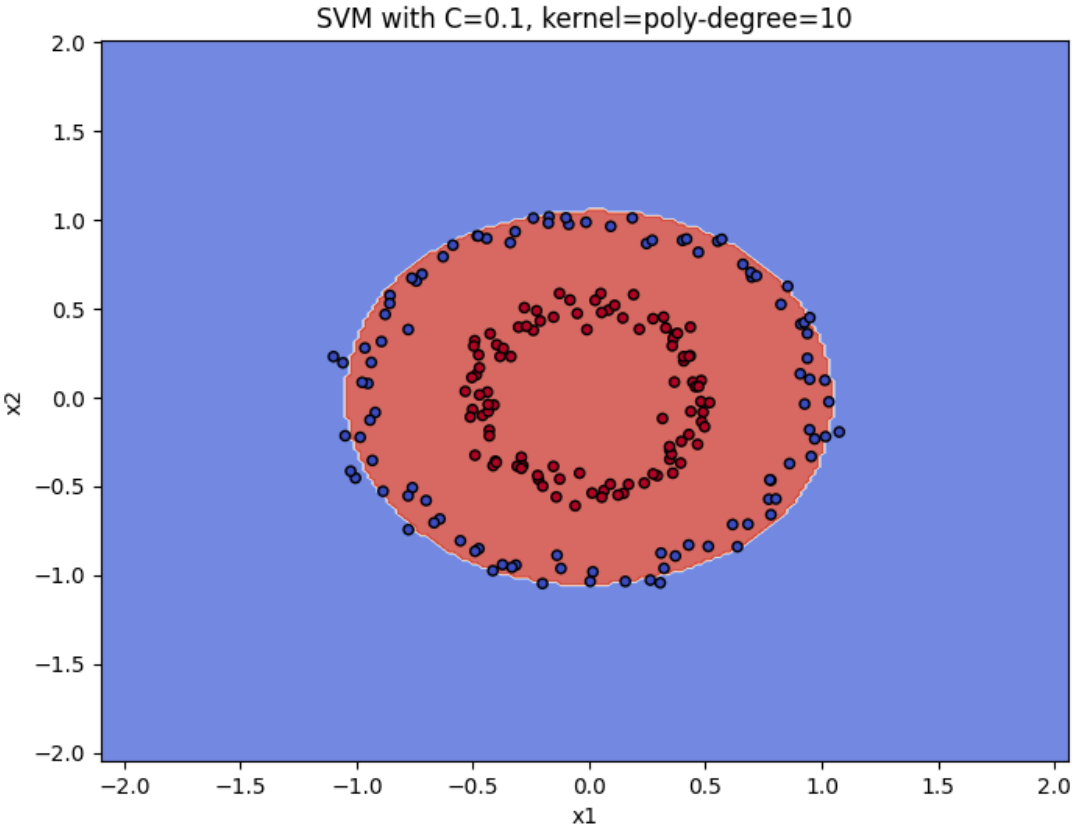
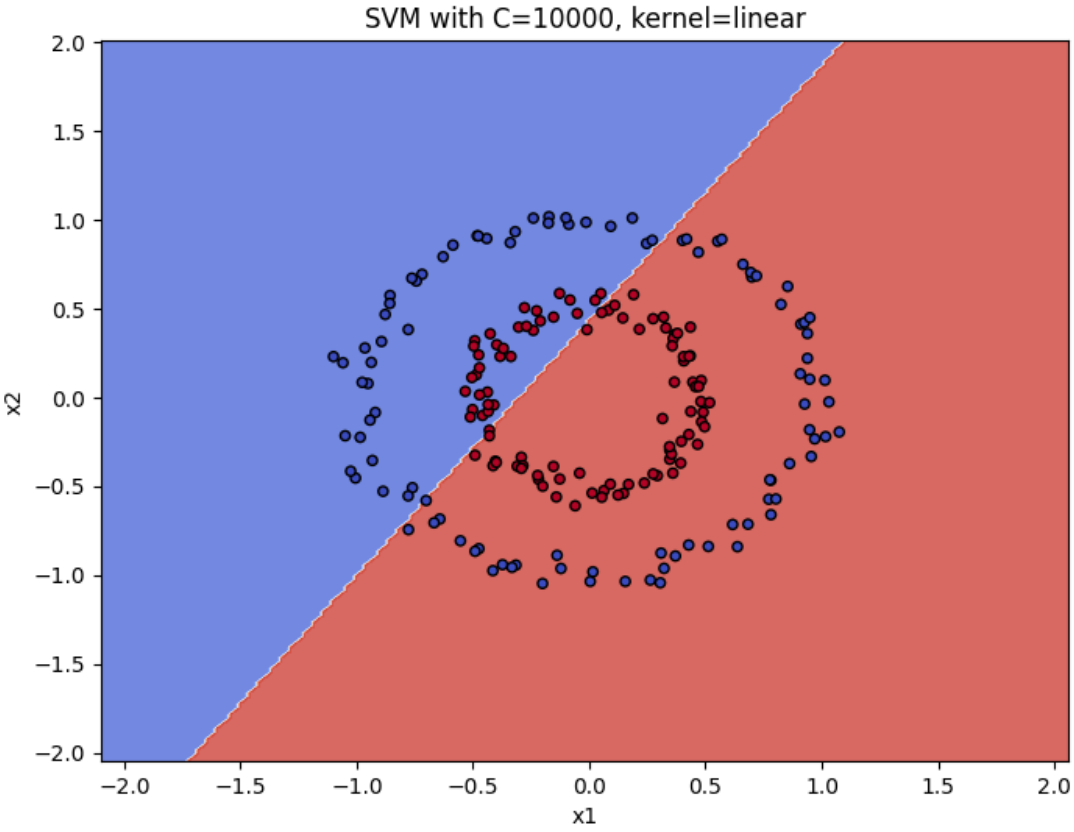
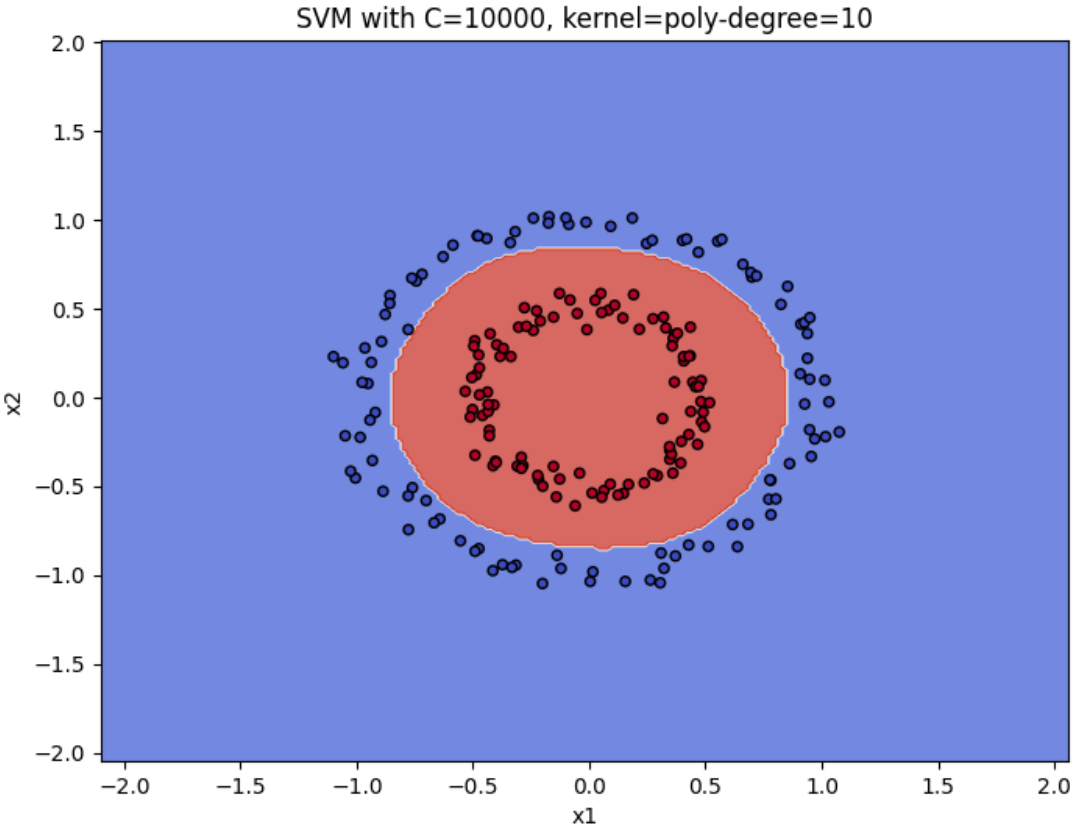
Part 2

Dataset 1




SVM with C=0.1, kernel=poly-degree=10
x2
x1
2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
−1.5
−2.0
−2.0
−1.5
−1.0
−0.5
0.0
0.5
1.0
1.5
2.0


SVM with C=10000, kernel=linear
2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
−1.5
−2.0
x2
−2.0
−1.5
−1.0
−0.5
0.0
0.5
1.0
1.5
2.0
x1


SVM with C=10000, kernel=poly-degree=10
2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
−1.5
−2.0
x2
−2.0
−1.5
−1.0
−0.5
0.0
0.5
1.0
1.5
2.0
x1

Dataset 2

	C	Kernel	Average Accuracy	Average Mean	Average STD
0	0.1	linear	0.877333	0.877333	0.082753
1	0.1	poly	0.633333	0.633333	0.085477
2	0.1	rbf	0.837333	0.837333	0.092528
3	0.1	sigmoid	0.830667	0.830667	0.083821
4	0.5	linear	0.953333	0.953333	0.055086
5	0.5	poly	0.662667	0.662667	0.083078
6	0.5	rbf	0.898667	0.898667	0.082928
7	0.5	sigmoid	0.872000	0.872000	0.053995
8	0.8	linear	0.954667	0.954667	0.049035
9	0.8	poly	0.670667	0.670667	0.078259
10	0.8	rbf	0.916000	0.916000	0.073816
11	0.8	sigmoid	0.830667	0.830667	0.071104
12	1.0	linear	0.954667	0.954667	0.049890
13	1.0	poly	0.665333	0.665333	0.076067
14	1.0	rbf	0.924000	0.924000	0.071269
15	1.0	sigmoid	0.821333	0.821333	0.076700
16	1.5	linear	0.954667	0.954667	0.047532
17	1.5	poly	0.672000	0.672000	0.082332
18	1.5	rbf	0.928000	0.928000	0.068498
19	1.5	sigmoid	0.798667	0.798667	0.084099
20	3.0	linear	0.958667	0.958667	0.046682
21	3.0	poly	0.682667	0.682667	0.072947
22	3.0	rbf	0.937333	0.937333	0.059241
23	3.0	sigmoid	0.772000	0.772000	0.082785
24	5.0	linear	0.953333	0.953333	0.049337
25	5.0	poly	0.701333	0.701333	0.076017
26	5.0	rbf	0.940000	0.940000	0.058620
27	5.0	sigmoid	0.774667	0.774667	0.088134
28	7.0	linear	0.953333	0.953333	0.047346
29	7.0	poly	0.701333	0.701333	0.078820
30	7.0	rbf	0.942667	0.942667	0.058600
31	7.0	sigmoid	0.773333	0.773333	0.086087
32	10.0	linear	0.954667	0.954667	0.050086
33	10.0	poly	0.697333	0.697333	0.081235
34	10.0	rbf	0.941333	0.941333	0.058545
35	10.0	sigmoid	0.776000	0.776000	0.091502

Part 3

Hyperparameter Search Results

KNN

	K	Average Accuracy	Average F1	Average Mean Accuracy	Average Mean F1	Average STD Accuracy	Average STD F1	CI Accuracy	CI F1
0	3	0.702575	0.799439	0.702575	0.799439	0.031596	0.022111	(0.69, 0.71)	(0.79, 0.81)
1	5	0.715064	0.812150	0.715064	0.812150	0.027342	0.018815	(0.7, 0.73)	(0.8, 0.82)
2	7	0.723922	0.821114	0.723922	0.821114	0.025828	0.017087	(0.71, 0.73)	(0.81, 0.83)
3	9	0.726383	0.824868	0.726383	0.824868	0.025195	0.016472	(0.72, 0.74)	(0.82, 0.83)

SVC

	C	Gamma	Kernel	Average Accuracy	Average F1	Average Mean Accuracy	Average Mean F1	Average STD Accuracy	Average STD F1	CI Accuracy	CI F1
0	1	scale	rbf	0.750966	0.838899	0.750966	0.838899	0.024442	0.015941	(0.74, 0.76)	(0.83, 0.85)
1	1	scale	linear	0.741657	0.822160	0.741657	0.822160	0.030642	0.022418	(0.73, 0.75)	(0.81, 0.83)
2	1	auto	rbf	0.751085	0.838916	0.751085	0.838916	0.024380	0.015912	(0.74, 0.76)	(0.83, 0.85)
3	1	auto	linear	0.741657	0.822160	0.741657	0.822160	0.030642	0.022418	(0.73, 0.75)	(0.81, 0.83)
4	10	scale	rbf	0.731866	0.814315	0.731866	0.814315	0.030343	0.022113	(0.72, 0.74)	(0.81, 0.82)
5	10	scale	linear	0.740717	0.821363	0.740717	0.821363	0.030760	0.022429	(0.73, 0.75)	(0.81, 0.83)
6	10	auto	rbf	0.731964	0.814141	0.731964	0.814141	0.030326	0.022259	(0.72, 0.74)	(0.81, 0.82)
7	10	auto	linear	0.740717	0.821363	0.740717	0.821363	0.030760	0.022429	(0.73, 0.75)	(0.81, 0.83)

Decision Tree

	Max Depth	Min Samples Split	Average Accuracy	Average F1	Average Mean Accuracy	Average Mean F1	Average STD Accuracy	Average STD F1	CI Accuracy	CI F1
0	5	4	0.692648	0.789036	0.692648	0.789036	0.032598	0.027482	(0.68, 0.71)	(0.78, 0.8)
1	5	6	0.692229	0.788803	0.692229	0.788803	0.032532	0.027578	(0.68, 0.7)	(0.78, 0.8)
2	7	4	0.685471	0.778488	0.685471	0.778488	0.035903	0.029865	(0.67, 0.7)	(0.77, 0.79)
3	7	6	0.685015	0.778105	0.685015	0.778105	0.035051	0.028470	(0.67, 0.7)	(0.77, 0.79)
4	9	4	0.676792	0.768240	0.676792	0.768240	0.037795	0.030729	(0.66, 0.69)	(0.76, 0.78)
5	9	6	0.675870	0.767431	0.675870	0.767431	0.037506	0.030549	(0.66, 0.69)	(0.76, 0.78)

Random Forest

	N Estimators	Max Depth	Min Samples Split	Average Accuracy	Average F1	Average Mean Accuracy	Average Mean F1	Average STD Accuracy	Average STD F1	CI Accuracy	CI F1
0	50	5	4	0.726211	0.832829	0.726211	0.832829	0.018239	0.010699	(0.72, 0.73)	(0.83, 0.84)
1	100	5	4	0.726032	0.833207	0.726032	0.833207	0.017178	0.009971	(0.72, 0.73)	(0.83, 0.84)
2	50	5	6	0.725832	0.832620	0.725832	0.832620	0.018603	0.011084	(0.72, 0.73)	(0.83, 0.84)
3	100	5	6	0.725514	0.832890	0.725514	0.832890	0.017372	0.010197	(0.72, 0.73)	(0.83, 0.84)
4	50	7	4	0.740081	0.836664	0.740081	0.836664	0.023201	0.014577	(0.73, 0.75)	(0.83, 0.84)
5	100	7	4	0.740995	0.837453	0.740995	0.837453	0.021535	0.013233	(0.73, 0.75)	(0.83, 0.84)
6	50	7	6	0.740076	0.836625	0.740076	0.836625	0.023125	0.014262	(0.73, 0.75)	(0.83, 0.84)
7	100	7	6	0.741596	0.838157	0.741596	0.838157	0.022894	0.014167	(0.73, 0.75)	(0.83, 0.84)

Final Evaluation Results

	Algorithm	Accuracy
0	KNN	0.750751
1	SVC	0.792793
2	Decision Tree	0.754491
3	Random Forest	0.777778

Decision Tree Training

The top 5 most important features from most important to less important are:
Attr1: Status of existing checking account
Attr2: Duration in month
Attr5: Credit amount
Attr3: Credit history
Attr6: Savings accounts/bonds

Support Vector Classifier Training

Here is the output of printing the support vectors(scaled back with StandardScaler()):

```
The support vectors are:
[[0.274      0.71243996 0.063      ... 0.404      1.15176175 0.037     ]
 [0.72000897 0.269      0.063      ... 0.404      1.15176175 0.037     ]
 [0.274      0.71243996 0.063      ... 0.404      1.15176175 0.037     ]
 ...
 [0.72000897 0.269      0.063      ... 0.404      1.15176175 0.037     ]
 [0.72000897 0.269      0.063      ... 0.89469746 1.15176175 0.037     ]
 [0.274      0.71243996 0.063      ... 0.404      1.15176175 0.037     ]]
```

Here is one of the support vectors:

```
[2.74000000e-01 7.12439962e-01 6.30000000e-02 3.94000000e-01
 5.99436610e+02 4.00000000e-02 4.90000000e-02 1.02909919e+00
 8.80000000e-02 2.93000000e-01 2.34000000e-01 1.03000000e-01
 1.81000000e-01 7.28998886e-01 1.20000000e-02 2.20000000e-02
 5.00000000e-02 0.00000000e+00 9.00000000e-03 9.70000000e-02
 1.20000000e-02 1.67929773e+07 1.09227600e+00 1.03000000e-01
 6.30000000e-02 4.80000000e-02 1.83000000e-01 6.20000000e-02
 1.72000000e-01 8.12369834e-01 1.74000000e-01 2.53000000e-01
 5.20931035e+00 5.00000000e-02 7.72493243e-01 5.48000000e-01
 9.20000000e-02 0.00000000e+00 1.19743244e+00 4.10000000e-02
 5.20000000e-02 5.05133180e+00 7.31973333e-01 2.32000000e-01
 3.32000000e-01 1.54000000e-01 2.85681147e+02 1.39000000e-01
 4.70000000e-02 1.20310667e+00 1.79000000e-01 1.16536158e+00
 1.08000000e-01 1.98436557e+00 2.20000000e-02 2.00000000e-01
 1.11280431e+00 1.48000000e-01 1.51690468e+00 1.08669746e+00
 4.04000000e-01 1.15176175e+00 3.70000000e-02]
```

I couldn't figure out how to extract feature importances or any other important information from this 😕