

CS 5433: Bigdata Management

Programming Assignment 3

Task 2 – Output

Group 4

Output Results for PartA:

prediction	Year	features
2019.6599680249049	2019.0	[0.011,0.0,0.919,...
2019.7068695042635	2020.0	[0.011,0.0,0.965,...
2016.9526883493056	2017.0	[0.016,0.0,0.877,...
2018.3079382199621	2018.0	[0.022,0.115,0.98...
2019.124822827811	2018.0	[0.038,0.0,0.856,...
2019.5989792076111	2021.0	[0.038,0.0,0.876,...
2019.6108093721991	2019.0	[0.038,0.0,0.888,...
2016.3784155412097	2017.0	[0.049,0.038,0.60...
2020.0829923571039	2021.0	[0.049,0.038,0.65...
2019.5629441635174	2019.0	[0.054,0.231,0.94...
2016.4310669178797	2017.0	[0.06,0.269,0.602...
2019.7920963222393	2019.0	[0.06,0.269,0.659...
2019.8593262626084	2021.0	[0.06,0.269,0.688...
2016.4429517220756	2017.0	[0.076,0.346,0.69...
2018.9371283132239	2018.0	[0.076,0.346,0.70...
2020.0737156980294	2019.0	[0.076,0.346,0.71...
2020.0973760272054	2020.0	[0.076,0.346,0.73...
2019.9178120161619	2020.0	[0.082,0.154,0.62...
2019.9796926800836	2019.0	[0.082,0.154,0.64...
2018.5560702047862	2018.0	[0.087,0.308,0.66...
2019.629987659562	2021.0	[0.087,0.308,0.71...
2018.6876945553227	2018.0	[0.092,0.038,0.76...
2019.8917530383087	2020.0	[0.092,0.038,0.87...
2018.2007661604232	2018.0	[0.098,0.038,0.79...
2019.8417642613454	2020.0	[0.098,0.038,0.84...
2017.3168362396666	2017.0	[0.103,0.231,0.89...
2018.4456122878844	2018.0	[0.103,0.231,0.92...
2019.5758057921732	2021.0	[0.103,0.231,0.93...
2019.5997619616849	2020.0	[0.103,0.231,0.96...
2017.1046111707344	2017.0	[0.109,0.154,0.69...
2020.1484769391539	2020.0	[0.109,0.154,0.81...
2018.9562888714165	2018.0	[0.12,0.154,0.801...
2020.060554275545	2019.0	[0.12,0.154,0.802...
2020.060554275545	2020.0	[0.12,0.154,0.802...
2019.8527542303623	2019.0	[0.125,0.077,0.64...
2017.3727435070696	2017.0	[0.13,0.462,0.633...
2019.8709857820218	2020.0	[0.13,0.462,0.814...
2019.8554403994342	2021.0	[0.141,0.769,0.69...

2019.8554403994342	2021.0	[0.141,0.769,0.69...
2020.2272015052897	2021.0	[0.147,0.385,0.76...
2020.2323532674434	2019.0	[0.147,0.385,0.77...
2020.2550027369552	2020.0	[0.147,0.385,0.79...
2020.2595439477984	2020.0	[0.152,0.654,0.80...
2018.4170124462992	2018.0	[0.158,0.0,0.604,...
2019.505160321904	2021.0	[0.158,0.0,0.647,...
2018.269757449315	2018.0	[0.163,0.538,0.66...
2019.7768929639376	2021.0	[0.163,0.538,0.69...
2019.7620095187888	2019.0	[0.163,0.538,0.7,...
2019.7994441074072	2020.0	[0.163,0.538,0.71...
2016.6032196770363	2017.0	[0.174,0.0,0.605,...
2018.3117742176364	2018.0	[0.174,0.0,0.686,...
2020.317189136891	2020.0	[0.174,0.0,0.756,...
2017.0722872382669	2017.0	[0.185,0.115,0.59...
2019.3622852796952	2019.0	[0.185,0.115,0.69...
2017.213540225763	2017.0	[0.19,0.538,0.557...
2018.3790161631598	2018.0	[0.196,0.192,0.67...
2020.1439677161513	2019.0	[0.196,0.192,0.7,...
2017.2292828432921	2017.0	[0.201,0.346,0.68...
2020.4294441402906	2019.0	[0.212,0.385,0.95...
2020.1250785650102	2019.0	[0.217,0.615,0.63...
2018.3345775499236	2018.0	[0.217,0.615,0.67...
2020.168626230091	2020.0	[0.223,0.615,0.73...
2017.5936011215204	2017.0	[0.223,0.962,0.56...
2020.3433839851793	2019.0	[0.228,0.0,0.798,...
2020.3621150791103	2020.0	[0.228,0.0,0.817,...
2019.6604684147796	2021.0	[0.234,0.0,0.879,...
2019.5134412652421	2021.0	[0.239,0.231,0.66...
2019.4790953163122	2019.0	[0.239,0.231,0.72...
2018.564732255747	2018.0	[0.239,0.231,0.74...
2018.1598944647476	2018.0	[0.245,0.192,0.78...
2020.081570131645	2021.0	[0.25,0.846,0.623...
2018.293401163434	2018.0	[0.255,0.154,0.63...
2020.1723372856	2021.0	[0.255,0.154,0.74...
2017.6443124210725	2017.0	[0.261,0.346,0.56...
2019.39244558066	2019.0	[0.261,0.346,0.61...
2019.4230801526057	2020.0	[0.261,0.346,0.64...
2018.461470032169	2018.0	[0.266,0.038,0.67...
2020.364114842615	2020.0	[0.266,0.038,0.74...
2020.1663219595391	2019.0	[0.272,0.192,0.61...

2020.1663219595391	2019.0	[0.272,0.192,0.61...
2020.1694718647375	2020.0	[0.283,0.077,0.88...
2020.1694718647375	2021.0	[0.283,0.077,0.88...
2020.0191695294468	2019.0	[0.288,0.077,0.73...
2018.461934551142	2018.0	[0.293,0.231,0.60...
2019.2976536823496	2018.0	[0.299,0.115,0.82...
2017.494406109878	2017.0	[0.304,0.423,0.59...
2019.364453398921	2019.0	[0.304,0.423,0.60...
2018.8826087784796	2018.0	[0.31,0.077,0.621...
2020.039699488902	2019.0	[0.31,0.077,0.635...
2019.0511006912238	2018.0	[0.326,0.077,0.89...
2020.1857420330562	2019.0	[0.326,0.077,0.90...
2017.6724929418676	2017.0	[0.332,0.885,0.6,...
2018.9210934965931	2018.0	[0.332,0.885,0.61...
2017.4631724863552	2017.0	[0.337,0.077,0.66...
2019.9621073481455	2021.0	[0.342,0.077,0.64...
2019.9325871351623	2020.0	[0.342,0.077,0.65...
2018.13612687311	2018.0	[0.348,0.038,0.66...
2020.293281504884	2021.0	[0.348,0.038,0.70...
2019.708864302203	2020.0	[0.353,0.0,0.812,...
2018.805166462856	2018.0	[0.364,0.038,0.69...
2018.8284011952387	2018.0	[0.37,0.038,0.739...
2019.9070480208886	2021.0	[0.37,0.038,0.753...

+-----+
only showing top 100 rows

```
2022-04-30 11:40:19,029 INFO datasources.FileSourceStrategy: Pruning directories with:
2022-04-30 11:40:19,029 INFO datasources.FileSourceStrategy: Pushed Filters:
2022-04-30 11:40:19,030 INFO datasources.FileSourceStrategy: Post-Scan Filters:
2022-04-30 11:40:19,030 INFO datasources.FileSourceStrategy: Output Data Schema: struct<INSTITUT
```

Output Results for PartB:

prediction	Year	features
2020.1276026413568	2019.0	[0.011,0.0,0.919,...
2020.203992290115	2020.0	[0.011,0.0,0.965,...
2017.2932738073287	2017.0	[0.016,0.0,0.877,...
2018.0033849272481	2018.0	[0.022,0.115,0.98...
2018.058098731962	2018.0	[0.038,0.0,0.856,...
2020.0168073610869	2021.0	[0.038,0.0,0.876,...
2020.0168073610869	2019.0	[0.038,0.0,0.888,...
2017.8982539682543	2017.0	[0.049,0.038,0.60...
2019.9220865825464	2021.0	[0.049,0.038,0.65...
2020.0016276403887	2019.0	[0.054,0.231,0.94...
2017.2745454545457	2017.0	[0.06,0.269,0.602...
2019.9931863827592	2019.0	[0.06,0.269,0.659...
2019.9517907367294	2021.0	[0.06,0.269,0.688...
2017.307568216942	2017.0	[0.076,0.346,0.69...
2018.0588073629103	2018.0	[0.076,0.346,0.70...
2020.0372479562998	2019.0	[0.076,0.346,0.71...
2020.1604505561693	2020.0	[0.076,0.346,0.73...
2019.9414770263909	2020.0	[0.082,0.154,0.62...
2020.1900975806977	2019.0	[0.082,0.154,0.64...
2018.0588073629103	2018.0	[0.087,0.308,0.66...
2019.83805965084	2021.0	[0.087,0.308,0.71...
2018.0836737924578	2018.0	[0.092,0.038,0.76...
2020.0407507845393	2020.0	[0.092,0.038,0.87...
2018.0295652629436	2018.0	[0.098,0.038,0.79...
2020.14525964841	2020.0	[0.098,0.038,0.84...
2017.2922637063189	2017.0	[0.103,0.231,0.89...
2018.0478293716926	2018.0	[0.103,0.231,0.92...
2019.8766633845219	2021.0	[0.103,0.231,0.93...
2019.6747709114036	2020.0	[0.103,0.231,0.96...
2017.1673214285715	2017.0	[0.109,0.154,0.69...
2020.1072549556307	2020.0	[0.109,0.154,0.81...
2018.0033849272481	2018.0	[0.12,0.154,0.801...
2020.1364658295647	2019.0	[0.12,0.154,0.802...
2020.1364658295647	2020.0	[0.12,0.154,0.802...
2020.01078711536	2019.0	[0.125,0.077,0.64...
2017.7877976190478	2017.0	[0.13,0.462,0.633...
2020.085874223584	2020.0	[0.13,0.462,0.814...
2019.9542207942886	2021.0	[0.141,0.769,0.69...

2017.1673214285715	2017.0	[0.109,0.154,0.69...
2020.1072549556307	2020.0	[0.109,0.154,0.81...
2018.0033849272481	2018.0	[0.12,0.154,0.801...
2020.1364658295647	2019.0	[0.12,0.154,0.802...
2020.1364658295647	2020.0	[0.12,0.154,0.802...
2020.01078711536	2019.0	[0.125,0.077,0.64...
2017.7877976190478	2017.0	[0.13,0.462,0.633...
2020.085874223584	2020.0	[0.13,0.462,0.814...
2019.9542207942886	2021.0	[0.141,0.769,0.69...
2020.1219962881023	2021.0	[0.147,0.385,0.76...
2020.211676576293	2019.0	[0.147,0.385,0.77...
2020.211676576293	2020.0	[0.147,0.385,0.79...
2020.197971607088	2020.0	[0.152,0.654,0.80...
2018.0185506939615	2018.0	[0.158,0.0,0.604,...
2020.0376238975143	2021.0	[0.158,0.0,0.647,...
2018.0470190948338	2018.0	[0.163,0.538,0.66...
2019.9190856591533	2021.0	[0.163,0.538,0.69...
2019.7713170737345	2019.0	[0.163,0.538,0.7,...
2019.9190856591533	2020.0	[0.163,0.538,0.71...
2017.4194444444445	2017.0	[0.174,0.0,0.605,...
2017.9643142180114	2018.0	[0.174,0.0,0.686,...
2020.0929569163145	2020.0	[0.174,0.0,0.756,...
2017.2045454545455	2017.0	[0.185,0.115,0.59...
2019.7157960279085	2019.0	[0.185,0.115,0.69...
2017.2	2017.0	[0.19,0.538,0.557...
2017.9696482377217	2018.0	[0.196,0.192,0.67...
2019.843842086851	2019.0	[0.196,0.192,0.7,...
2017.161929271709	2017.0	[0.201,0.346,0.68...
2020.2321952551451	2019.0	[0.212,0.385,0.95...
2019.8916510080915	2019.0	[0.217,0.615,0.63...
2018.052353114544	2018.0	[0.217,0.615,0.67...
2020.1874763067717	2020.0	[0.223,0.615,0.73...
2017.0	2017.0	[0.223,0.962,0.56...
2020.3563330946695	2019.0	[0.228,0.0,0.798,...
2020.3190315073675	2020.0	[0.228,0.0,0.817,...
2020.0008642763937	2021.0	[0.234,0.0,0.879,...
2019.8685422274543	2021.0	[0.239,0.231,0.66...
2019.8833366002764	2019.0	[0.239,0.231,0.72...
2018.1512283030813	2018.0	[0.239,0.231,0.74...
2018.1310060808592	2018.0	[0.245,0.192,0.78...
2019.8755607421867	2021.0	[0.25,0.846,0.623...

2019.8755607421867	2021.0	[0.25,0.846,0.623...
2018.0665943447552	2018.0	[0.255,0.154,0.63...
2020.0840683497997	2021.0	[0.255,0.154,0.74...
2017.0	2017.0	[0.261,0.346,0.56...
2019.9199084914314	2019.0	[0.261,0.346,0.61...
2019.9505840595314	2020.0	[0.261,0.346,0.64...
2017.9962037932776	2018.0	[0.266,0.038,0.67...
2020.0358721016037	2020.0	[0.266,0.038,0.74...
2019.9213271513302	2019.0	[0.272,0.192,0.61...
2019.6383194659952	2020.0	[0.283,0.077,0.88...
2019.6383194659952	2021.0	[0.283,0.077,0.88...
2020.045367149362	2019.0	[0.288,0.077,0.73...
2018.1692729161837	2018.0	[0.293,0.231,0.60...
2018.1310060808592	2018.0	[0.299,0.115,0.82...
2017.2045454545455	2017.0	[0.304,0.423,0.59...
2019.3221831129704	2019.0	[0.304,0.423,0.60...
2018.0707610114216	2018.0	[0.31,0.077,0.621...
2020.0075323502983	2019.0	[0.31,0.077,0.635...
2018.111860072792	2018.0	[0.326,0.077,0.89...
2020.0978580072583	2019.0	[0.326,0.077,0.90...
2017.309902597403	2017.0	[0.332,0.885,0.6,...
2018.3367291847578	2018.0	[0.332,0.885,0.61...
2017.1355436507936	2017.0	[0.337,0.077,0.66...
2020.0022112317843	2021.0	[0.342,0.077,0.64...
2019.7284385416824	2020.0	[0.342,0.077,0.65...
2017.9962037932776	2018.0	[0.348,0.038,0.66...
2019.775411008557	2021.0	[0.348,0.038,0.70...
2020.0694607901191	2020.0	[0.353,0.0,0.812,...
2018.0449783030813	2018.0	[0.364,0.038,0.69...
2018.108230884678	2018.0	[0.37,0.038,0.739...
2019.915939148387	2021.0	[0.37,0.038,0.753...
2017.9036517040622	2018.0	[0.375,0.231,0.59...
2020.1018348736573	2021.0	[0.375,0.231,0.62...
2020.006807100513	2021.0	[0.38,0.115,0.792...
2017.147876984127	2017.0	[0.386,0.577,0.66...
2018.3899736235223	2018.0	[0.391,0.346,0.60...
2019.789052984713	2021.0	[0.391,0.346,0.61...
2020.0547411455732	2019.0	[0.413,0.385,0.62...
2017.0	2017.0	[0.424,0.0,0.572,...
2018.279461061702	2020.0	[0.435,0.192,0.65...
2018.4938808800584	2019.0	[0.446,0.0,0.639,...

2018.4938808800584	2019.0	[0.446,0.0,0.639,...
2017.1938769841267	2017.0	[0.446,0.0,0.663,...
2019.7261489290136	2021.0	[0.457,0.077,0.65...
2018.4422770496355	2019.0	[0.467,0.423,0.66...
2020.1000129074425	2020.0	[0.478,0.038,0.70...
2018.3396433159276	2018.0	[0.484,0.731,0.63...
2018.27850346473	2018.0	[0.489,0.808,0.72...
2020.2948935542179	2020.0	[0.505,0.423,0.63...
2020.2079879963853	2020.0	[0.516,0.0,0.665,...
2018.0795890081376	2018.0	[0.522,0.0,0.604,...
2020.2423193429502	2020.0	[0.527,0.192,0.75...
2020.2639161351126	2020.0	[0.538,0.308,0.65...
2019.7074502015607	2021.0	[0.543,0.0,0.676,...
2020.0334280517932	2019.0	[0.549,0.0,0.607,...
2017.0	2017.0	[0.56,0.077,0.0,0...
2018.4343602919512	2019.0	[0.576,0.308,0.69...
2018.1198456770865	2018.0	[0.576,0.308,0.71...
2018.3789785657013	2019.0	[0.582,0.923,0.60...
2020.3882325469772	2019.0	[0.592,0.0,0.737,...
2018.8582749749694	2019.0	[0.598,0.192,0.62...
2018.851722549883	2019.0	[0.609,0.077,0.61...
2017.7536309523814	2017.0	[0.63,0.115,0.613...
2018.7225485759457	2021.0	[0.674,0.115,0.66...
2017.9563095238095	2017.0	[0.701,0.731,0.61...
2017.05	2017.0	[0.707,0.192,0.55...
2017.05	2017.0	[0.712,0.692,0.58...
2017.7052857142858	2017.0	[0.717,0.308,0.62...
2017.1386389452334	2017.0	[0.723,0.154,0.79...
2019.9053667440135	2021.0	[0.755,0.077,0.60...
2017.1764575163402	2017.0	[0.766,0.192,0.64...
2017.0	2017.0	[0.777,0.0,0.577,...
2017.0	2017.0	[0.804,0.269,0.57...
2019.0748700045174	2019.0	[0.81,0.077,0.614...
2017.0	2017.0	[0.821,0.077,0.55...
2018.3449913331272	2018.0	[0.826,1.0,0.597,...
2018.792587206855	2021.0	[0.848,0.577,0.64...
2019.7208886491487	2021.0	[0.886,0.0,0.778,...
2017.1588611674556	2017.0	[0.891,0.0,0.679,...
2017.0	2017.0	[0.908,0.192,0.56...
2019.9894303612189	2020.0	[0.913,0.0,0.798,...
2017.2	2017.0	[0.967,0.115,0.55...

2019.9894303612189	2020.0	[0.913,0.0,0.798,...]
2017.2	2017.0	[0.967,0.115,0.55...]
2017.75	2017.0	[0.978,0.0,0.582,...]
2017.8	2017.0	[0.984,0.231,0.55...]

+-----+-----+-----+

2022-04-30 11:39:11,958 INFO datasources.FileSourceStrategy: Pruning directories with:
2022-04-30 11:39:11,960 INFO datasources.FileSourceStrategy: Pushed Filters:
2022-04-30 11:39:11,960 INFO datasources.FileSourceStrategy: Post-Scan Filters:
2022-04-30 11:39:11,960 INFO datasources.FileSourceStrategy: Output Data Schema: struct
. 1 more field>
2022-04-30 11:39:12,024 INFO output.FileOutputCommitter: File Output Committer Algorith
2022-04-30 11:39:12,024 INFO output.FileOutputCommitter: FileOutputCommitter skip clear