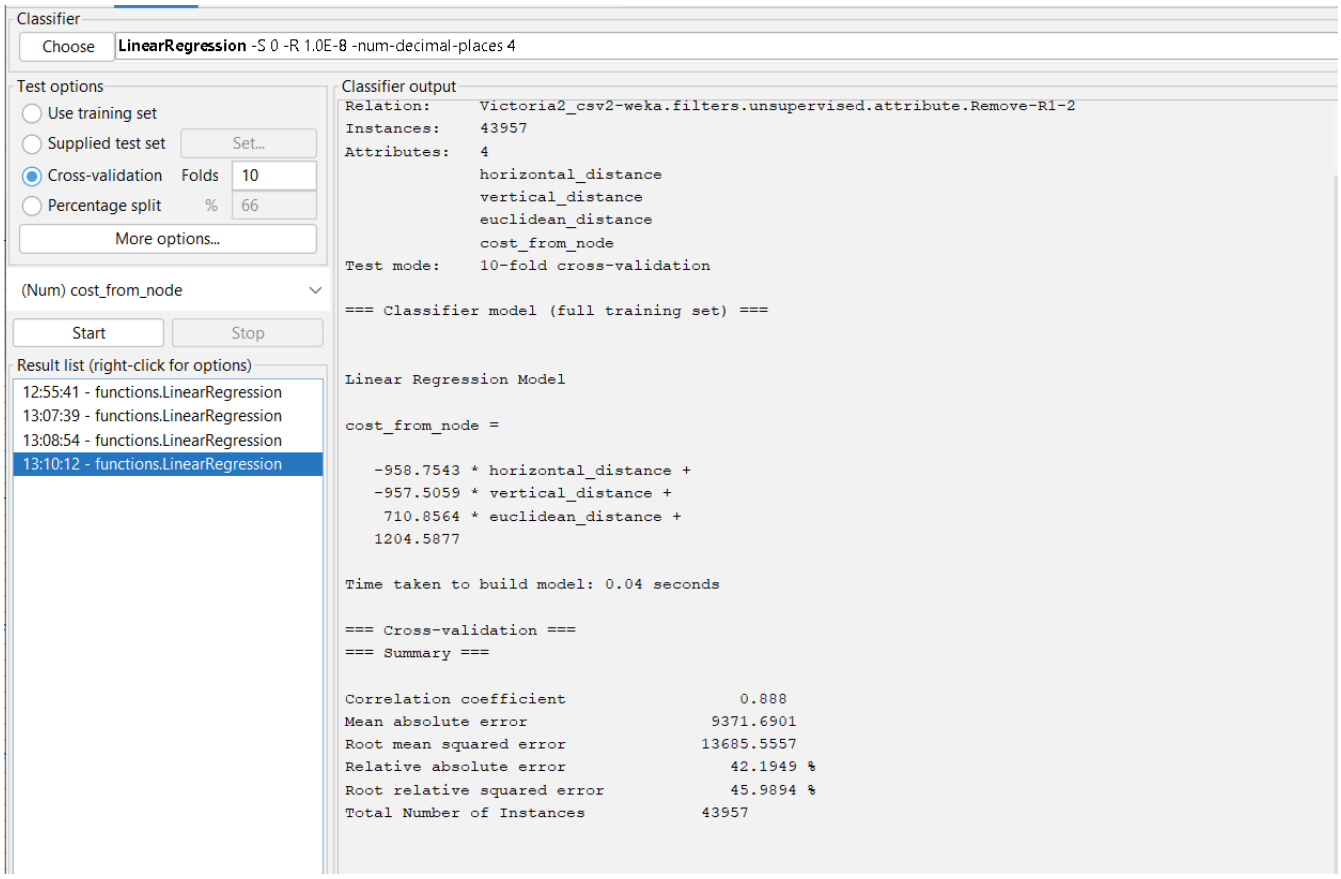
Chandler Garcia, Victoria Grasso, and Dillon McCarthy
CS4341 Assignment 3: Machine Learning

For this assignment, we set off to create a new heuristic for our Astar algorithm using machine learning and the code we created in assignment 1. Using Victoria's Assignment 1 code as a base, we gathered data we thought might be necessary and useful to train our model based on the specifications given in the assignment. The features we selected to base our model on were the horizontal distance, vertical distance, and euclidean distance, taken at each point along the path in respect to the current position and the goal. In our CSV file (titled "Data.csv"), we also saved the number of moves taken, the cost from node, and the number of nodes expanded. We decided to train our model using a linear regression model using Weka after gathering data from a 190 by 190 board for a couple hours, which consisted of a total of around 44,000 data points. With this data and the linear regression model, we were given the equation for heuristic 7:

$$cost\ from\ node\ =\ -\ 958.7543\ *\ horizontal\ distance\ -\ 957.5059\ *\ vertical\ distance + 710.8564\ *\ euclidean\ distance\ +\ 1204.5877$$

The degree of model fit using Weka has a correlation coefficient of 0.888, a relative absolute error of 42%, and a root relative error of 46%.

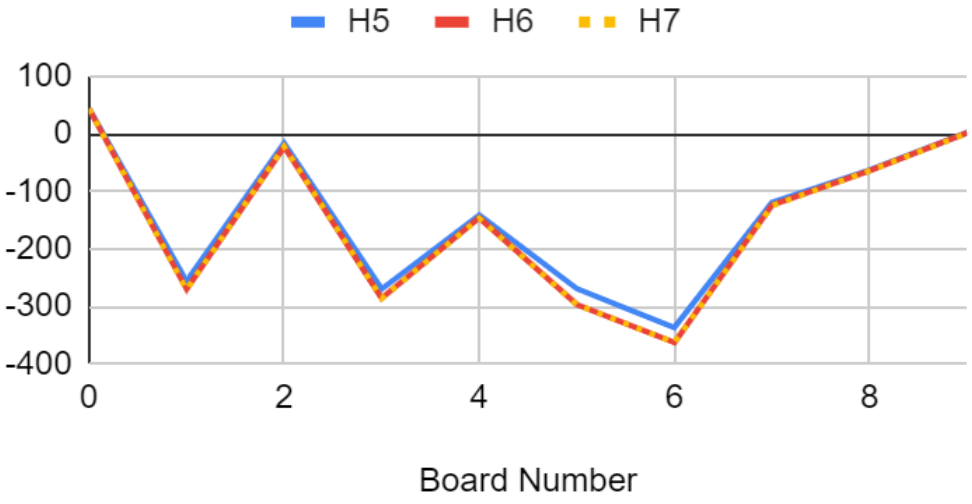


To collect data and compare the heuristics, we ran the code base on 10 boards of size 190x190 and we got the following results:

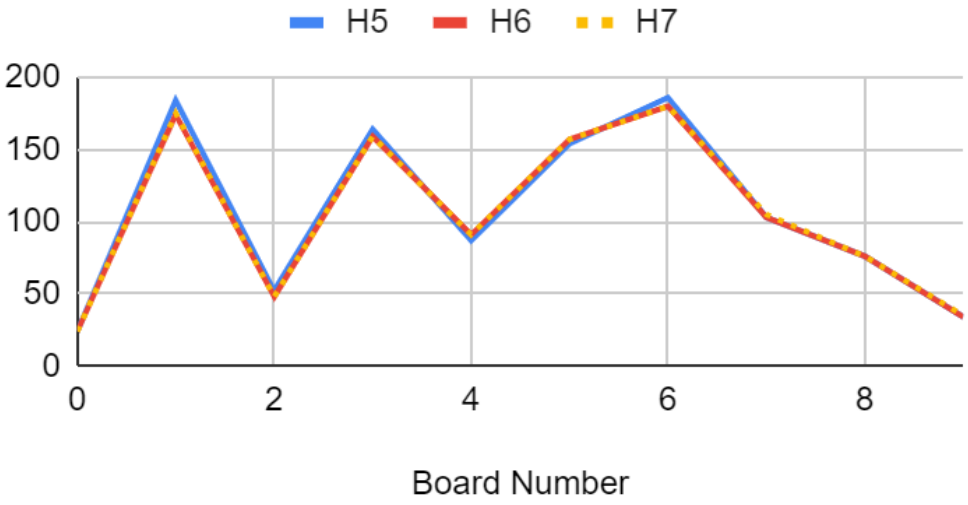
Heuristic 5	Scores (solution cost)	Moves	Nodes Expanded
0	44	24	2678
1	-256	184	179072
2	-16	52	17925
3	-270	164	117066
4	-141	87	49531
5	-269	154	167178
6	-337	186	153511
7	-119	103	56864
8	-63	76	32214
9	2	34	11349
Heuristic 6	44	24	52
	-270	174	918
	-22	48	201
	-286	159	857
	-146	91	4478
	-297	157	734
	-363	180	1309
	-124	103	706
	-64	76	638
	2	34	442
Heurstic 7	44	24	63
	-271	175	1008

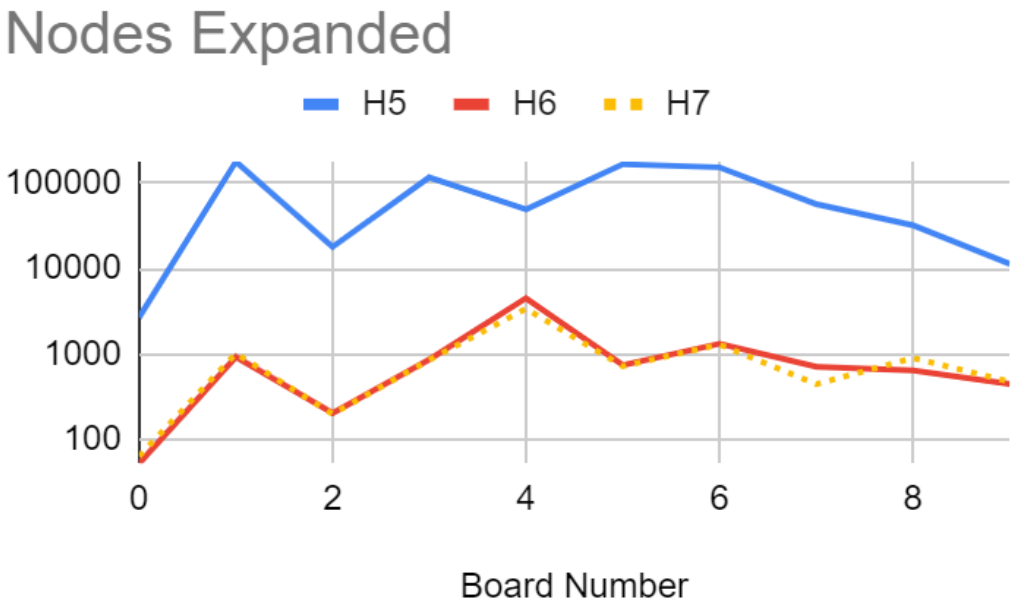
	-22	48	197
	-286	159	847
	-147	91	3376
	-297	157	716
	-363	180	1286
	-125	105	441
	-64	76	893
	1	35	472

Scores



Moves





We computed the following for heuristics 5,6, and 7.

Heuristic	Average # Nodes Expanded	Branching Factor	Average Time	Average Memory	Average Score
5	78738.8	1.113346555	5.830731797	0.0064149504 GB	-142.5
6	1033.5	1.068335645	0.06318976879	0.0001193984 GB	-152.6
7	929.9	1.067261447	0.09134047031	8.54016e-05 GB	-153

Our learned heuristic (7) has a smaller average number of nodes expanded from the 10 boards than heuristics 5 and 6 and therefore also has a smaller branching factor. However heuristic 7 has a greater negative score than the other two heuristics. This means that while heuristic 7 looks at fewer nodes, it finds a slightly less optimal solution on average compared to heuristics 5 and 6.

Appendix A

Board/Heuristic	Score	Moves	Nodes Expanded	Memory (GB)	Time (sec)	Branching Factor
Board 0						
5	44	24	2678	0.001081344	0.178463459	1.389394069
6	44	24	52	0.00012288	0.004654169083	1.178962901
7	44	24	63	0.00012288	0.006979465485	1.188427035
Board 1						
5	-256	184	179072	0.065011712	10.69504333	1.067945435
6	-270	174	918	0.000712704	0.05584812164	1.039986811
7	-271	175	1008	0.000765952	0.09045648575	1.040309661
Board 2						
5	-16	52	17925	0.001519616	1.036774397	1.207250216
6	-22	48	201	4.10E-06	0.0130546093	1.116820176
7	-22	48	197	0	0.01894903183	1.116352578
Board 3						
5	-270	164	117066	0.01026048	7.791581631	1.07375467
6	-286	159	857	0.000114688	0.05837726593	1.0433894
7	-286	159	847	0.000114688	0.0844783783	1.043312381
Board 4						
5	-141	87	49531	0.00487424	3.391590118	1.13230677

						2
6	-146	91	4478	0.0012288	0.2460868359	1.09678576
7	-147	91	3376	0.000516096	0.3547165394	1.093386364
Board 5						
5	-269	154	167178	0.026607616	13.95030284	1.081226665
6	-297	157	734	0.000114688	0.05210638046	1.042924431
7	-297	157	716	8.19E-05	0.07024264336	1.04275951
Board 6						
5	-337	186	153511	0.010338304	12.88116646	1.066307516
6	-363	180	1309	0.000118784	0.08178067207	1.040677899
7	-363	180	1286	4.92E-05	0.1090638638	1.040575415
Board 7						
5	-119	105	56864	0.004870144	4.939192295	1.109900803
6	-124	103	706	6.96E-05	0.04586100578	1.065757256
7	-125	105	441	5.32E-05	0.04886889458	1.059705356
Board 8						
5	-63	76	32214	0.002326528	2.590668917	1.146347735
6	-64	76	638	9.83E-05	0.04522275925	1.088693262
7	-64	76	893	4.10E-06	0.07380080223	1.093520661

Board 9						
5	2	34	11349	0.001409024	0.8525345325	1.316022937
6	2	34	442	0	0.02890586853	1.196207545
7	1	35	472	0	0.05584859848	1.192335138