

## CSC3022H ML ASSIGNMENT REPORT

Discount factor = 0.5

A lower discount factor reduces the “bouncing” effect when running the program as it does not value spots where the mines used to be.

Learning rate = 1

The learning rate is set to 1 to place more value on new experiences.

Mine reward = 100

The value is positive to encourage the collection of mines.

Super mine Reward = -100

The value is -100 so as to punish collision with a super mine. The sweeper will negatively learn to avoid them.

Movement reward = 0

The encourage exploration, the movement reward is set to 0 instead of a negative value

The table below is obtained after 50 iterations of the 3 environments.

Table

Sweepers	Mines	SuperMines	AverageMines	MaxMines	Averagedeaths
1	30	0	9.94	15	0
1	25	5	7.92	16	0.4
1	5	25	0.2	2	1