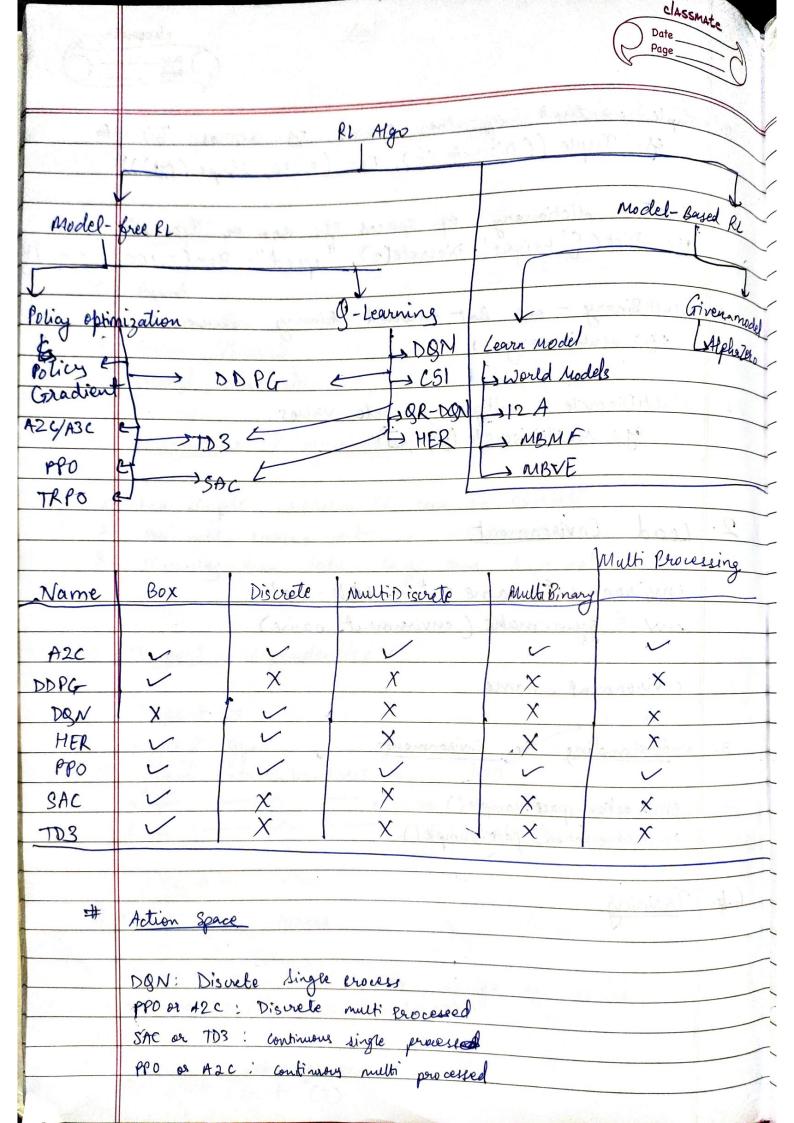
(RL focuses on teaching agents through Trial of
-	Concepts:
	Agent
2. E	arironnent
3.	Action Valored
4.	Reward, Observation
	Limitations & considerations
	g constillations
1	For simple peroblem RL can be overled
2	Thumes envis a med in markerian
3.	Training con take long time & is not always stable
	I save a is not already stable
1.	Import dependencies
	import os
	import gym journ stable baselines 3 import PPO prom
	1 MPert Dimmi Vector
	from - cralvation import evaluate-po
<u>#</u>	Open ti gym spaces:
-11	
1.	Box - n dimentional tensor, range of values eg: Box (0,1, shape = (3,3))
	eg: Box (0,1, shape = (3,3))
2.	Discrete - set of items

			Classmat Date Page	•
7.			Dafe Page	- 9
	Tuple tuple of		4	
	Tuple tuple of other space eg: Tuple ((Discrete (2), Box Dict dictionar	eg. 1307	cor Discre	fe
- 4	Box	(0,100, st	rape (1,))))	
6.	Dict - diction		· ix	
	es: Dict (5' lain of spaces	eg: 800 0	r discrete	31.4.3
	Discrete(2),	"speed":	Box (0,100,))
5.	Dict - dictionary of spaces eg: Dict ({ 'height': Discrete(2), Multi Binary - one hot encoded' eg: MultiBinary(4)			با ا
in close	ee one hot encoded!	binerry ra	lues without	Telias ob-
	5. Mult Binary (4)	0	1	
	Multipo	3 -390	0 + 1	POUR !
	MultiDiscrete - multiple discrete	values		
	MultiDiscrete - multiple discrete eg: MultiDiscrete ([5,2,2])	3	2012	1 CAV SA
	JV8AD /	1	tet vision	039
			13 PM	OPATS
2.	Load Environment			
2717	in the first of the first of			
	environment_name = t cartroli	e - vo	x x x 2	Name
	en = gyn. make (environment.	name)		•
340.00				NOA
	environment name	X		
	y y		V 1	31133
0	understanding the Environment		1	100 g
3.	understanding the entrance			
	Dr. 0.000 40.000 ()	~		
	env. action-space.sample()			745
	env. observation_space.sample().	1 /		200
4.6.	Training.		i i	
	V		Holise Se	42
1	talen.	M	de man	
.,	2000 Assistance of House.	minus Sign	3 Sk 1	
	The state of the state of the state of the state of	are dealer 2	di .	
	harden og diener in	Service .	73 133	
			LA COLLEGE TO	



Lan	Understaing Training Métrics:
	Tot A2C Algo
1.	Evaluation
	Ep-len-mean, ep-raw mean
2-	Time Metrics:
Torus	Eps, iterations, time_elapsed, total_timesteps
- (-1)	en alegant TXII
3.	Loss Metrics:
	Entropy loss, Policy loss, value loss
4.	Other Metrics:
	Explained varionce; Learning rate, nupdates
(3)	problem of Makoki -
	100 Am into vit «- 1
5	Sare Model & Reload model
	to save legal?
6.	Evaluation
	The state of the s
7.	Test Model.
-	
G.	Viening Wogs in Tensor Board.
	Company with the state of the s
	Training Staratergies:
1-	Train for ronger
3.	Ryperperameter tuning
3	Try diff algorithm.
9.	Addion a Calll II do
10-	Adding a Callback to traing stage
11.	Changing policies. I stage Vsing Alternate Algo.
	There is the state of the state

MIP COUNT