

کیوں ڈریں زندگی میں کیا ہو گا  
چھ نہ ہو گا تو تجربہ ہو گا۔

جاوید اختر



MUSIC  
Make Unification  
Simple in  
Image Classification



# Project Members

(in no particular order)

**Abdul Wahab**



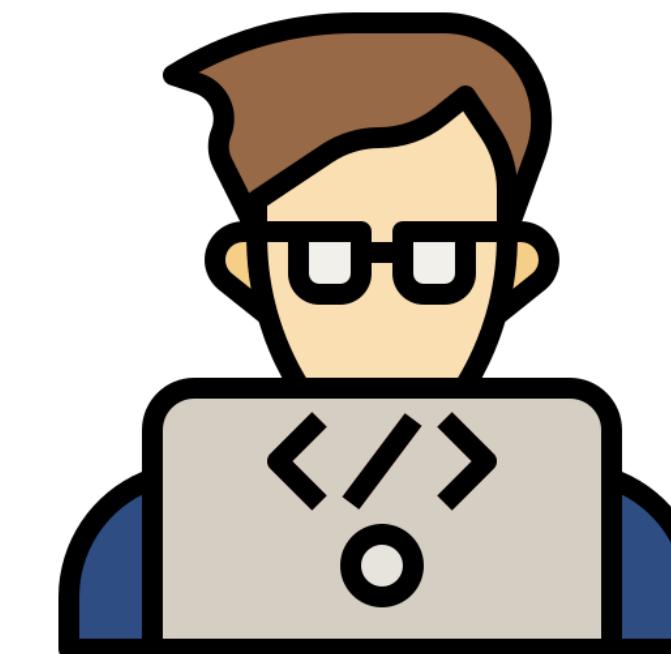
**Muhammad Shahbaz**



**Abdullah Asghar**



**Hamza Waheed**



# Availability of pre-trained models

# Availability of pre-trained models

Vision

0.24	0.55	0.27	0.47	0.07	0.73	0.69	0.8	0.56	0.98
0.88	0.31	0.46	0.8	0.16	0.95	0.09	0.61	0.22	0.75
0.85	0.06	0.9	0.95	0.44	0.53	0.02	0.18	0.76	0.77
0.01	0.2	0.35	0.32	0.29	0.56	0.07	0.67	0.92	0.95
0.93	0.86	0.23	0.28	0.61	0.19	0.06	0.84	0.35	0.16
0.63	0.88	0.18	0.73	0.73	0.3	0.07	0.53	0.47	0.07
0.36	0.08	0.52	0.29	0.95	0.1	0.06	0.41	0.62	0.88
0.2	0.97	0.6	0.81	0.78	0.06	0.71	0.61	0.35	0.22
0.28	0.25	0.03	0.75	0.54	0.27	0.54	0.98	0.74	0.89
0.39	0.3	0.75	0.21	0.53	0.72	0.33	0.94	0.04	0.45

# Availability of pre-trained models

Vision

0.54	0.88	0.87	0.34	0.29	0.85	0.36	0.41	0.11	0.41
0.36	0.62	0.28	0.24	0.14	0.93	0.79	0.37	0.63	0.19
0.04	0.75	0.32	0.5	0.71	0.52	0.46	0.21	0.04	0.64
0.65	0.45	0.48	0.59	0.31	0.5	0.63	0.11	0.54	0.23
0.93	0.22	0.56	0.8	0.61	0.08	0.2	0.37	0.69	0.8
0.64	0.35	0.33	0.58	0.13	0.84	0.68	0.17	0.49	0.75
0.04	0.83	0.77	0.93	0.06	0.54	0.17	0.88	0.89	0.22
0.06	0.28	0.4	0.53	0.02	0.92	0.17	0.81	0.72	0.17
0.83	0.28	0.08	0.38	0.39	0.32	0.75	0.17	0.58	0.73
0.51	0.18	0.58	0.26	0.99	0.15	0.6	0.31	0.49	0.36

Speech

0.42	0.77	0.4	0.42	0.48	0.33	0.93	0.67	0.99	0.32
0.76	0.6	0.42	0.75	0.12	0.26	0.59	0.3	0.91	0.42
0.49	0.48	0.81	0.09	0.15	0.57	0.49	0.46	0.36	0.55
0.95	0.86	0.29	0.16	0.8	0.03	0.43	0.15	0.81	0.59
0.46	0.54	0.01	0.4	0.08	0.32	0.97	0.17	0.58	0.73
0.28	0.33	0.05	0.74	0.45	0.02	0.1	0.87	0.52	0.09
0.67	0.91	0.67	0.32	0.9	0.54	0.75	0.91	0.52	0.82
0.87	0.43	0.92	0.48	0.32	0.01	0.54	0.55	0.81	0.59
0.12	0.02	0.83	0.06	0.58	0.37	0.37	0.72	0.55	0.16
0.67	0.45	0.31	0.26	0.67	0.42	0.19	0.29	0.06	1

# Availability of pre-trained models

Vision

0.86	0.68	0.89	0.96	0.38	0.11	0.46	0.45	0.36	0.65
0.18	0.4	0.72	0.33	0.74	0.38	0.36	0.6	0.47	0.78
0.69	0.54	0.95	0.19	0.49	0.01	0.34	0.22	0.98	0.29
0.76	0.95	0.17	0.22	0.83	0.21	0.1	0.45	0.85	0.42
0.63	0.33	0.5	0.49	0.74	0.18	0.72	0.75	0.65	0.34
0.66	0.88	0.71	0.53	0.73	0.93	0.66	0.94	0.96	0.59
0.89	0.47	0.74	0.17	0.3	0.21	0.73	0.3	0.38	0.74
0.58	0.99	0.22	0.25	0.98	0.68	0.76	0.54	0.49	0.17
0.14	0.24	0.35	0.61	0.55	0.82	0.93	0.97	0.29	0.24
0.27	0.5	0.1	0.2	0.86	0.4	0.43	0.35	0.92	0.21

NLP

0.24	0.23	0.53	0.39	0.33	0.5	0.81	0.05	0.08	0.54
0.81	0.45	0.9	0.68	0.66	0.73	0.73	0.9	0.37	0.71
0.44	0.11	0.22	0.17	0.18	0.76	0.3	0.08	0.25	0.84
0.16	0.5	0.91	0.36	0.7	0.83	0.72	0.05	0.81	0.5
0.83	0.71	0.79	0.35	0.34	0.93	0.43	0.41	0.75	0.42
0.16	0.22	0.67	0.43	0.12	0.83	0.96	0.45	0.48	0.74
0.66	0.04	0.45	0.63	0.85	0.43	0.86	0.15	0.24	0.72
0.12	0.38	0.19	0.94	0.39	0.83	0.85	0.58	0.51	0.99
0.91	0.01	0.8	0.43	0.58	0.47	0.15	0.56	0.14	0.89
0.95	0.85	0.54	0.66	0.74	0.8	0.02	0.74	0.68	0.46

Speech

0.68	0.33	0.8	0.12	0.33	0.32	0.83	0.43	0.49	0.78
0.22	0.41	0.71	0.59	0.98	0.24	0.2	0.21	0.28	0.91
0.5	0.14	0.51	0.83	0.25	0.13	0.65	0.26	0.11	0.89
0.92	0.36	0.95	0.47	0.06	0.7	0.4	0.71	0.4	0.32
0.23	0.18	0.97	0.31	0.35	0.62	0.73	0.4	0.58	0.38
0.79	0.53	0.88	0.1	0.61	0.07	0.79	0.65	0.18	0.39
0.93	0.5	0.72	0.31	0.25	0.48	0.11	0.7	0.72	0.67
0.54	0.56	0.57	0.44	0.18	0.55	0.05	0.58	0.86	0.59
0.66	0.49	0.63	0.86	0.31	0.42	0.7	0.7	0.64	0.75
0.2	0.31	0.98	0.62	0.54	0.99	0.65	0.5	0.39	0.17





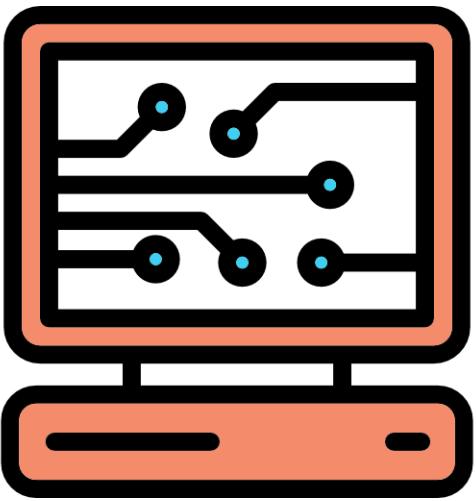




**BUT we can't always retrain them**

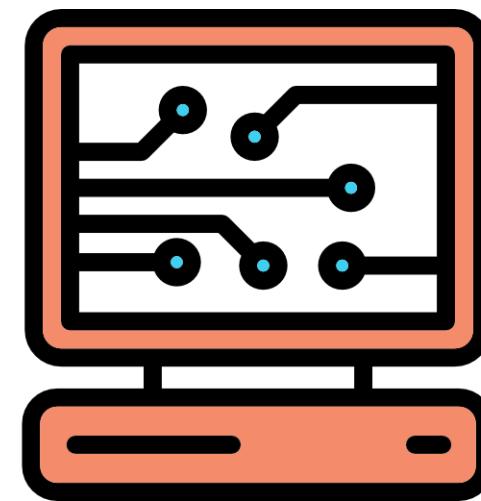
**BUT** we can't always retrain them

Resource intensive

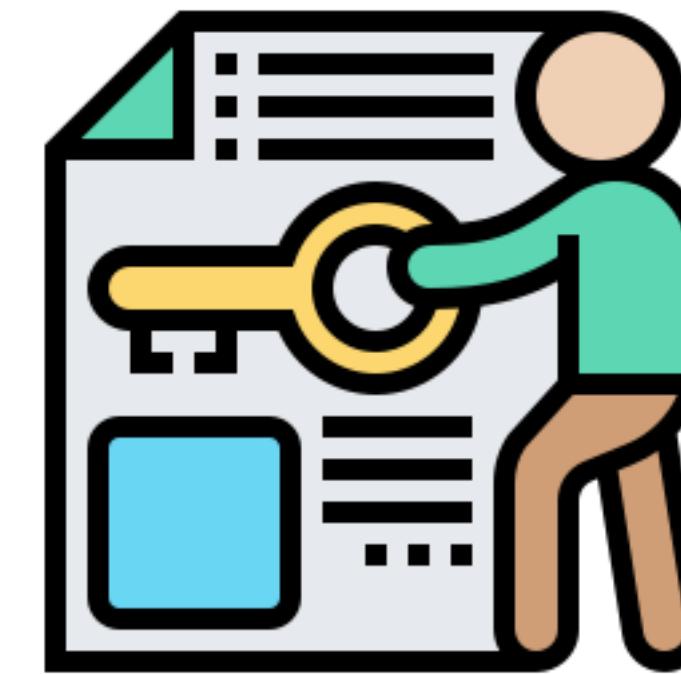


**BUT** we can't always retrain them

Resource intensive

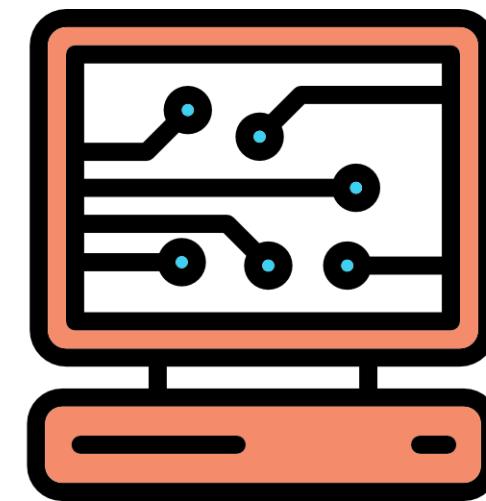


Data ownership issues

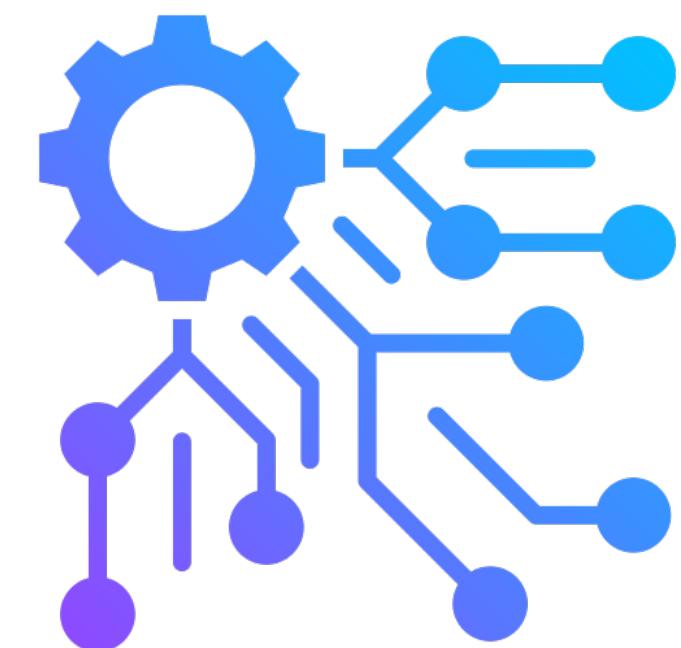


**BUT** we can't always retrain them

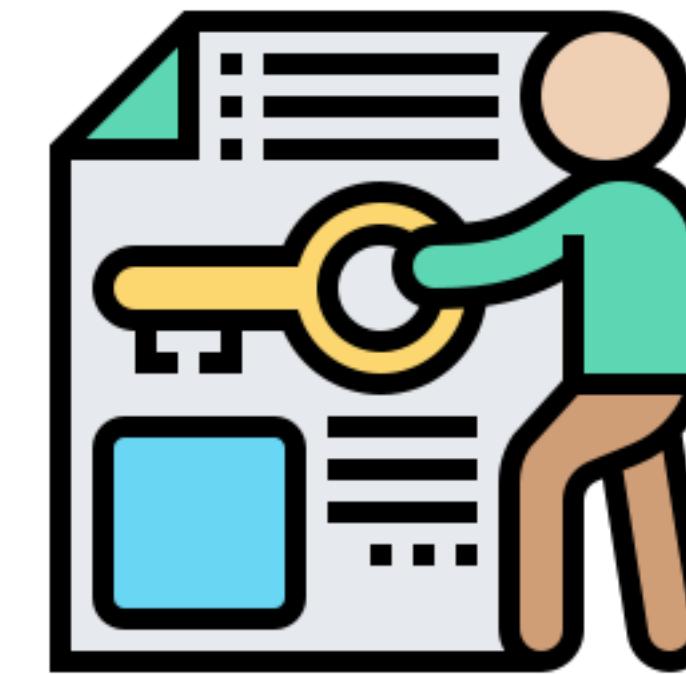
Resource intensive



Complexity

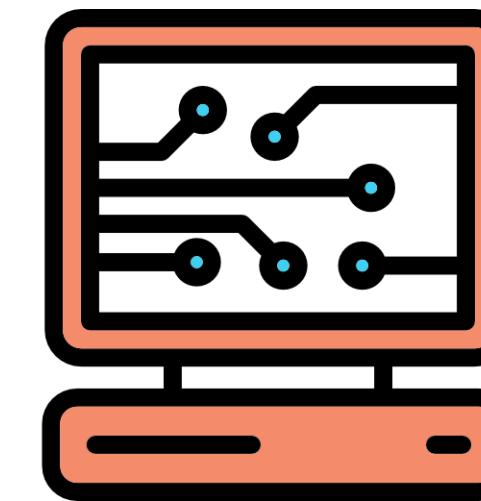


Data ownership issues

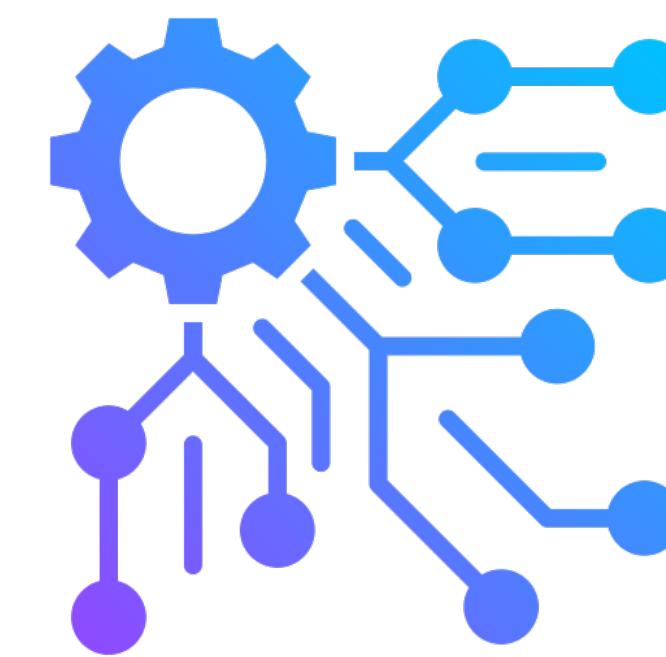


# BUT we can't always retrain them

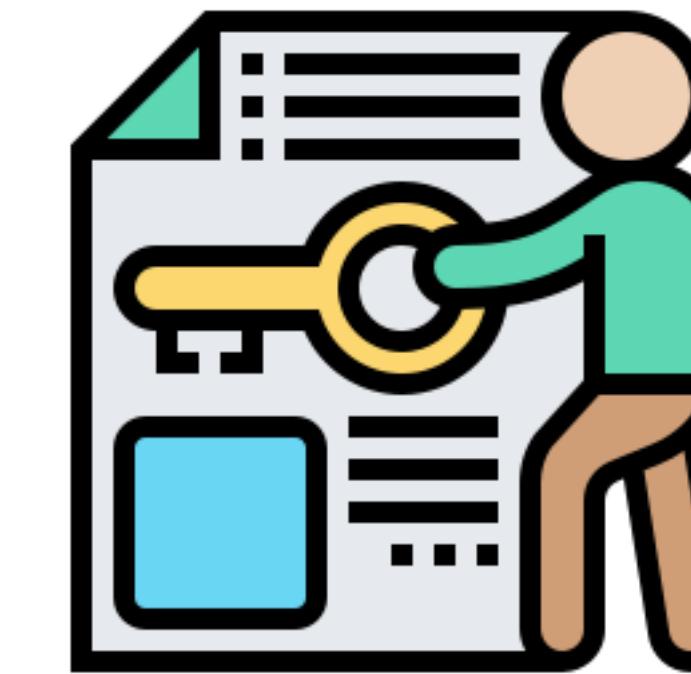
Resource intensive



Complexity



Data ownership issues



Unavailability of original training data



**We fine-tune them for specific tasks**

# We fine-tune them for specific tasks

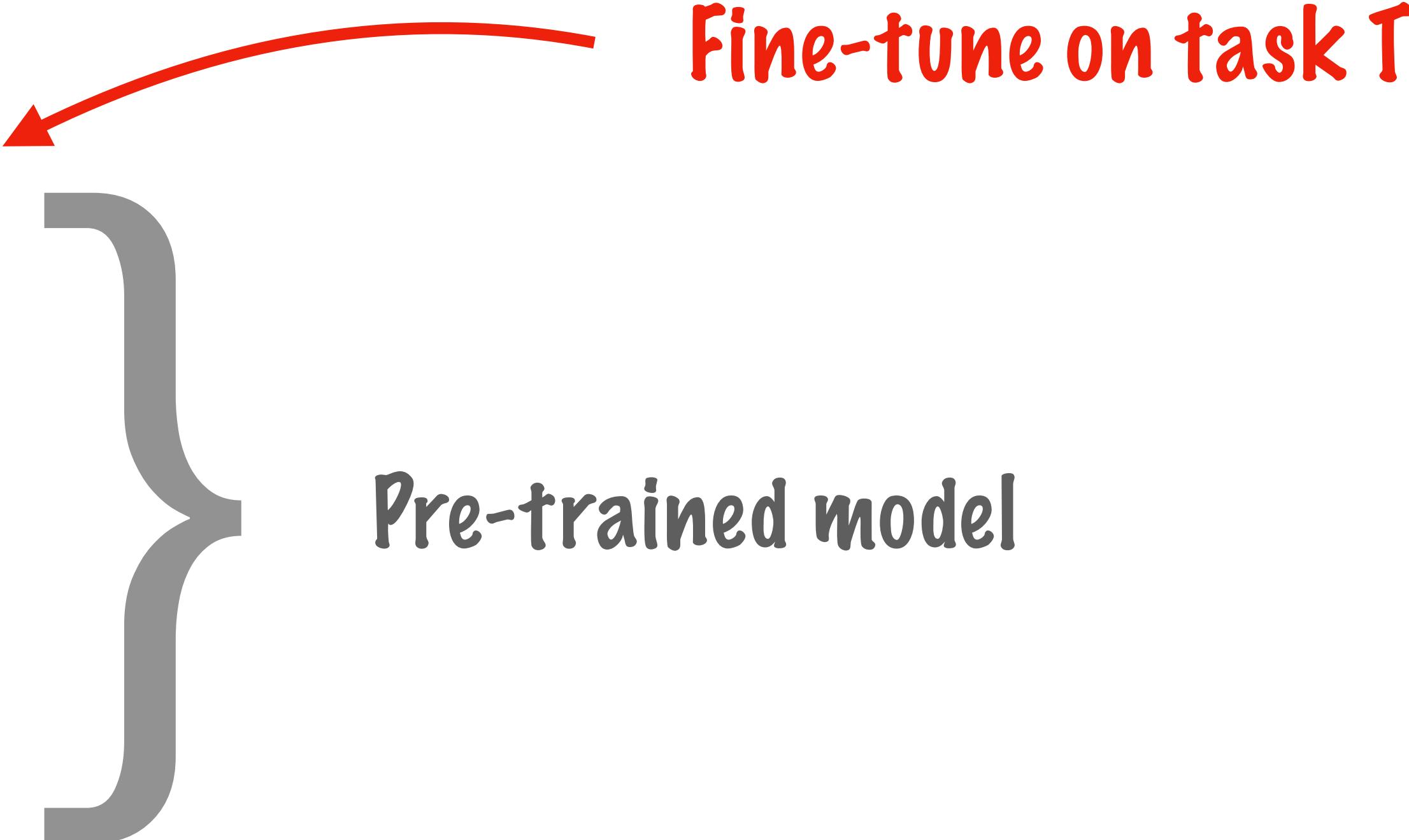
0.49	0.49	0.61	0.11	0.49	0.92	0.25	0.38	0.89	0.71
0.57	0.03	0.06	0.93	0.52	0.9	0.27	0.37	0.5	0.07
0.52	0.11	0.55	0.89	0.42	0.71	0.43	0.08	0.42	0.42
0.51	0.53	0.98	0.66	0.4	0.47	0.55	0.58	0.28	0.32
0.91	0.67	0.23	0.55	0.9	0.51	0.78	0.57	0.2	0.22
0.35	0.03	0.24	0.4	0.24	0.78	0.03	0.75	0.63	0.4
0.24	0.04	0.63	0.15	0.61	0.09	0.98	0.29	0.64	0.06
0.32	0.38	0.21	0.8	0.29	0.99	0.35	0.04	0.72	0.16
0.77	0.56	0.47	0.92	0.17	0.58	0.13	0.01	0.62	0.97
0.61	0.72	0.46	0.1	0.02	1	0.58	0.79	0.04	0.79



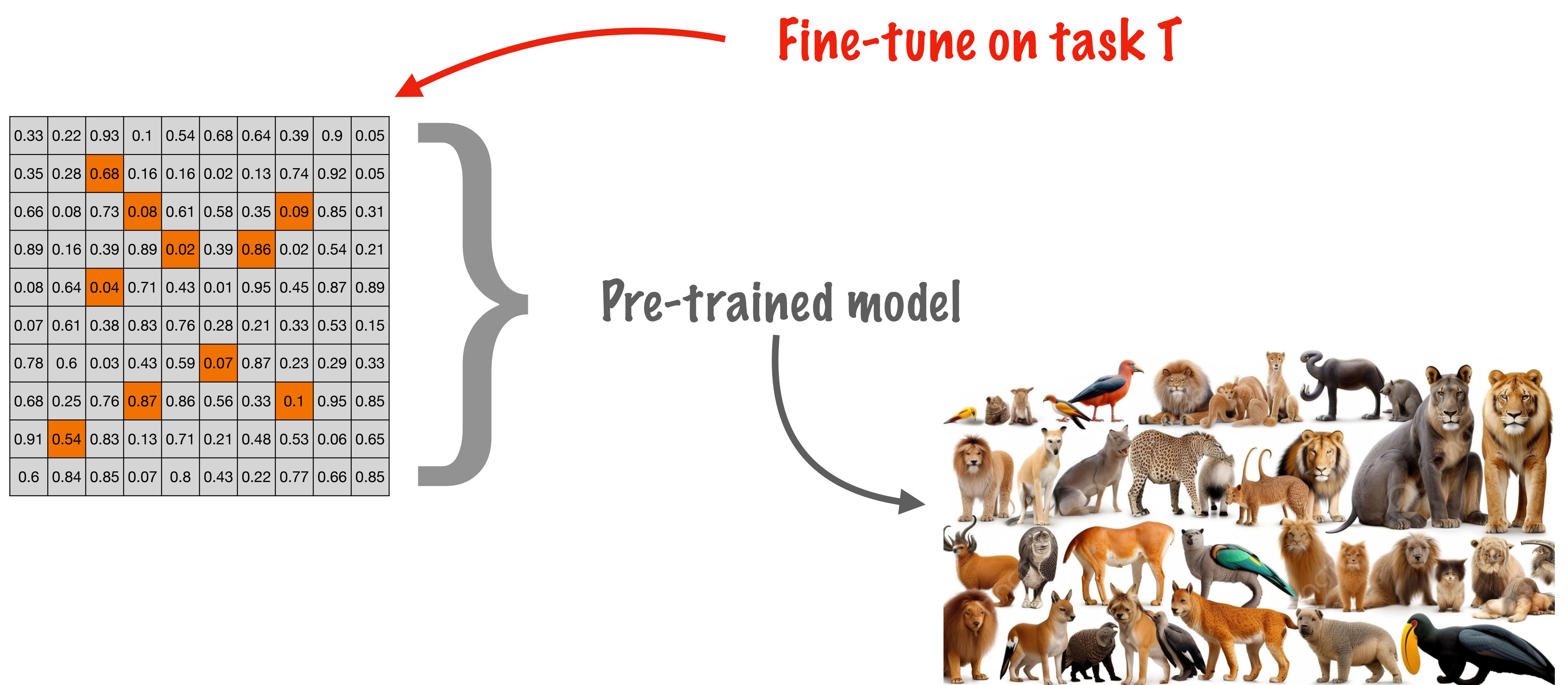
Pre-trained model

# We fine-tune them for specific tasks

0.1	0.54	0.51	0.25	0.52	0.55	0.23	0.02	0.73	0.2
0.16	0.76	0.33	0.7	0.99	0.57	0.93	0.57	0.36	0.52
0.55	0.31	0.6	0.59	0.57	0.67	0.83	0.42	0.82	0.74
0.18	0.25	0.05	0.75	0.87	0.9	0.74	0.06	0.75	0.61
0.38	0.62	0.98	0.01	0.48	0.15	0.31	0.42	0.09	0.71
0.05	0.47	0.65	0.72	0.86	0.78	0.38	0.65	0.15	0.06
0.8	0.76	0.96	0.95	0.81	0.94	0.6	0.67	0.52	0.49
0.2	0.94	0.13	0.1	0.53	0.37	0.28	0.14	0.8	0.79
0.54	0.55	0.7	0.51	0.47	0.12	0.22	0.51	0.39	0.46
0.49	0.27	0.95	0.52	0.26	0.29	0.76	0.31	0.61	0.5



# We fine-tune them for specific tasks



# We fine-tune them for specific tasks

0.01	0.57	0.92	0.56	0.01	0.18	0.77	0.93	0.28	0.34
0.06	0.32	0.95	0.58	0.77	0.28	0.67	0.83	0.41	0.73
0.82	0.98	0.9	0.15	0.28	0.2	0.82	0.45	0.81	0.65
0.56	0.72	0.65	0.15	0.05	0.34	0.56	0.9	0.99	0.09
0.25	0.32	0.63	0.61	0.48	1	0.27	0.22	0.36	0.13
0.84	0.18	0.34	0.8	0.81	0.24	0.76	0.79	0.01	0.06
0.07	0.67	0.53	0.75	0.56	0.46	0.62	0.98	0.3	0.08
0.98	0.68	0.4	0.93	0.63	0.62	0.68	0.02	0.34	0.7
0.68	0.79	0.39	0.47	0.84	0.82	0.31	0.35	0.61	0
0.94	0.52	0.68	0.32	0.78	0.77	0.68	0.99	0.21	0.85



# What is a Task Vector?

# What is a Task Vector?

Fine-tuned model

0.77	0.57	0.11	0.23	0.93	0.34	0.22	0.36	0.85	0.24
0.15	0.77	0.95	0.95	0.62	0.87	0.53	0.46	0.02	0.14
0.65	0.38	0.33	0.57	0.16	0.81	0.24	0.87	0.8	0.59
0.58	0.59	0.81	0.2	0.65	0.69	0.18	0.86	0.3	0.16
0.39	0.13	0.49	0.24	0.35	0.57	0.61	0.83	0.23	0.13
0.97	0.94	0.75	0.62	0.72	0.26	0.45	0.27	0.37	0.25
0.6	0.97	0.19	0.08	0.19	0.94	0.86	0.36	0.52	0.26
0.71	0.99	0.49	0.78	0.02	0.74	0.32	0.94	0.24	0.81
0.49	0.55	0.62	0.3	0.4	0.9	0.73	0.2	0.8	0.02
0.61	0.35	0.02	0.69	0.46	0.58	0.54	0.12	0.61	0.93

# What is a Task Vector?

**Fine-tuned model**

0.66	0.36	0.64	0.75	0.21	0.66	0.93	0.18	0.68	0.22	
0.67	0.82	0.04	0.85	0.44	0.25	0.98	0.36	0.84	0.11	
0.72	0.46	0.52	0.85	0.38	0.42	0.97	0.72	0.94	0.17	
0.22	0.54	0.33	0.79	0.59	0.24	0.02	0	0.95	0.94	
0.24	0.01	0.52	0.33	0.47	0.11	0.25	0.59	0.78	0.85	
0.77	0.75	0.23	0.88	0.91	0.28	0.19	0.05	0.76	0.13	
0.47	0.48	0.03	0.99	0.05	0.36	0.1	0.74	0.87	0.63	
0.48	0.85	0.92	0.46	0.79	0.71	0.3	0.85	0.63	0.04	
0.88	0.21	0.83	0.22	0.99	0.31	0.06	1	0.25	0.58	
0.97	0.15	0.11	0.05	0.24	0.7	0.52	0.86	0.67	0.18	

**Pre-trained model**

0.02	0.29	0.87	0.47	0.63	0.55	0.79	0.1	0.38	0.05	
0.92	0.02	0.76	0.16	0.79	0	0.63	0.14	0.76	0.74	
0.98	0.6	0.34	0.66	0.15	0.86	0.64	0.75	0.57	0.17	
0.98	0.91	0.82	0.51	0.62	0.89	0.72	0.59	0.1	0.6	
0.86	0.81	0.56	0.15	0.54	0.96	0.57	0.37	0.83	0.83	
0.41	0.42	0.03	0.78	0.88	0.16	0.18	0.11	0.91	0.06	
0.8	0.31	0.92	0.67	0.05	0.82	0.05	0.28	0.87	0.81	
0.99	0.71	0.92	0.27	0.96	0.11	0.34	0.83	0.98	0.39	
0.52	0.65	0.78	0.39	0.44	0.1	0.39	0.08	0.13	0.75	
0.25	0.95	0.94	0.6	0.83	0.55	0.07	0.43	0.49	0.85	

# What is a Task Vector?

**Fine-tuned model**

0.01	0.3	0.5	0.63	0.21	0.16	0.14	0.13	0.46	0.26
0.84	0.24	0.22	0.26	0.27	0.07	0.43	0.33	0.14	0.04
0.75	0.18	0.77	0	0.36	0.73	0.89	0.44	0.61	0.95
0.46	0.87	0.11	0.36	0.69	0.53	0.72	0.41	0.12	0.45
0.93	0.09	0.45	0.74	0	0.64	0.48	0.5	0.19	0.53
0.89	0.96	0.14	0.48	0.62	0.12	0.99	0.61	0.04	0.09
0.71	0.39	0.51	0.24	0.88	0.49	0.89	0.54	0.31	0.18
0.99	0.56	0.72	0.08	0.52	0.51	0.96	0.23	0.9	0.72
0.34	0.21	0.67	0.06	0.27	0.52	0.5	0.87	0.62	0.25
0.52	0.73	0.44	0.1	0.43	0.74	0.67	0.83	0.01	0.85

**Pre-trained model**

0.16	0.51	0.78	0.09	0.9	0.12	0.04	0.32	0.25	0.83
0.32	0.59	0.25	0.34	0.53	0.85	0.94	0.98	0.52	0.83
0.35	0.66	0.52	0.45	0.68	0.21	0.02	0.2	0.81	0.64
0.7	0.1	0.43	0.96	0.31	0.96	0.44	0.13	0.89	0.98
0.67	0.9	0.95	0.01	0.36	0.7	0.88	0	0.8	0.61
0.55	0.68	0.78	0.91	0.03	0.75	0.24	0.27	0.16	0.12
0.73	0.89	0.43	0.6	0.17	0.07	0.97	0.24	0.19	0.67
0.88	0.6	0.3	0.62	0.72	0.05	0.48	0.04	0.94	0.83
0.01	0.82	0.77	0.23	0.17	0.3	0.59	0.25	0.54	0.44
0.37	0.2	0.3	0.25	0.83	0.16	0.12	0.64	0.71	0.69

# What is a Task Vector?

# Fine-tuned model

0.46	0.03	0.07	0.87	0.36	0.76	0.29	0.98	0.9	0.17
0.24	0.03	0.24	0.98	0.49	0.55	0.93	0.38	0.72	0.82
0.9	0.6	0.48	0.88	0.39	0.32	0.44	0.39	0.76	0.96
0.84	0.15	0.42	0.72	0.55	0.79	0.13	0.34	0.38	0.91
0.92	0.98	0.94	0.98	0.05	0.38	0.22	0.65	0.94	0.85
0.03	0.26	0.44	0.3	0.77	0.86	0.99	0.32	0.49	0.54
0.8	0.04	0.76	0.26	0.56	0.33	0.46	0.19	0.02	0.78
0.94	0.19	0.06	0.71	0.79	0.05	0.77	0.75	0.38	0.26
0.31	0.69	0.31	0.17	0.25	0.54	0.6	0.52	0.3	0.15
0.07	0.4	0.33	0.03	0.88	0.47	1	0.09	0.56	0.87

# Pre-trained model

0.15	0.65	0.72	0.19	0.26	0.73	0.33	0.19	0.03	0.72
0.78	0.25	0.8	0.41	0.75	0.74	0.79	0.52	0.67	0.84
0.57	0.08	0.02	0.41	0.56	0.15	0.2	0.36	0.96	0.21
0.46	0.44	0.81	0.94	0.18	0.43	0.94	0.37	0.49	0.15
0.05	0.75	0.63	0.83	0.23	0.57	0.84	0.73	0.1	0.68
0.2	0.26	0.04	0.38	0.58	0.53	0.91	0.56	0.79	0.25
0.55	0.37	0.47	0.63	0.18	0.33	0.97	0.47	0.17	1
0.85	0.16	0.06	0.39	0.42	0.81	0.64	0.08	0.84	0.88
0.42	0.14	0.73	0.38	0.7	0.15	0.28	0.89	0.68	0.37
0.91	0.72	0.88	0.77	0.23	0.51	0.9	0.88	0.6	0.81

# A Task Vector











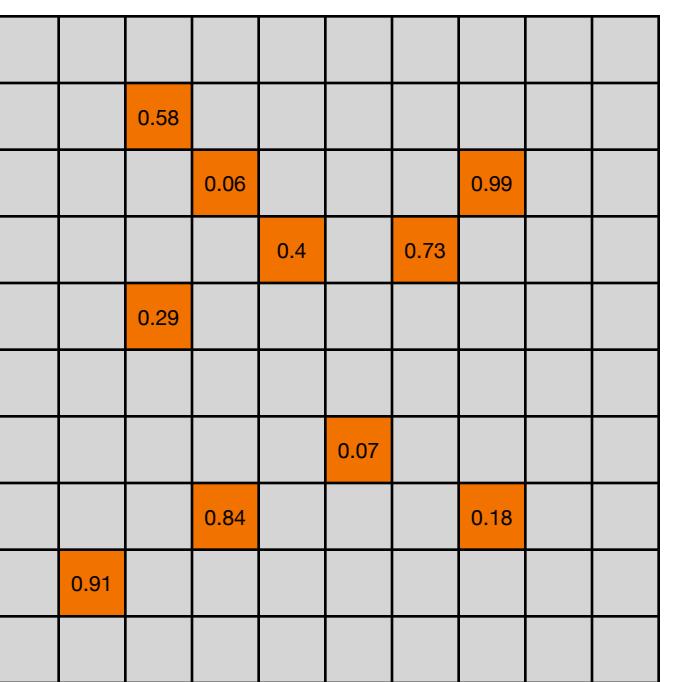
# Our Initial Hypothesis

*Having three tasks  $T_1$ ,  $T_2$  and  $T_3$*

# Our Initial Hypothesis

*Having three tasks  $T_1$ ,  $T_2$  and  $T_3$*

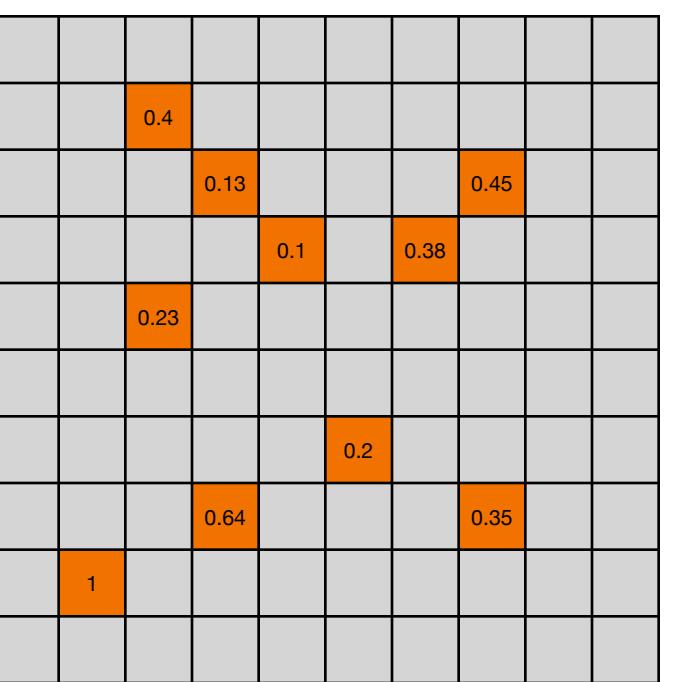
**tv1**



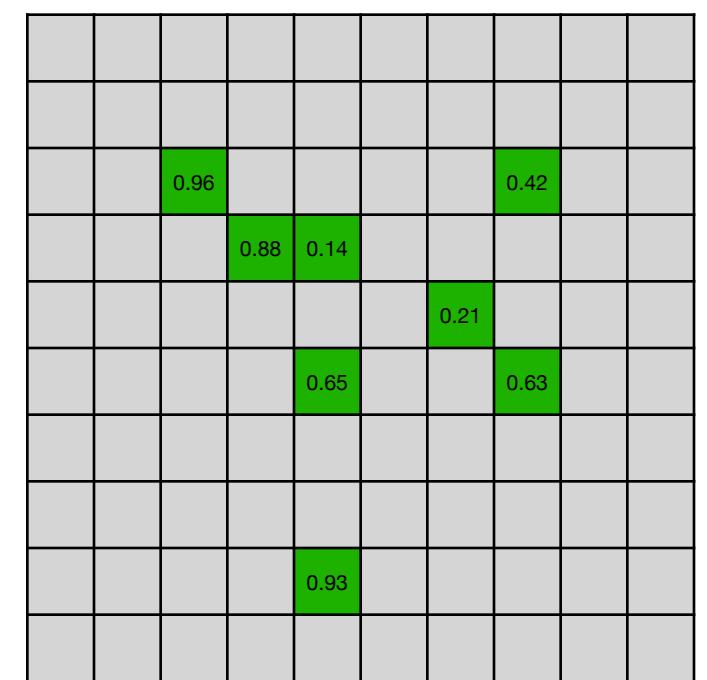
# Our Initial Hypothesis

*Having three tasks  $T_1$ ,  $T_2$  and  $T_3$*

**tv1**



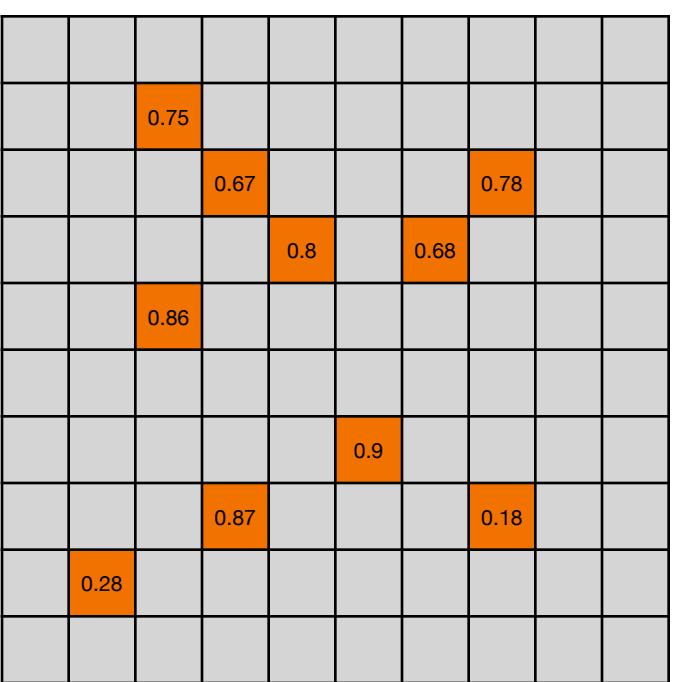
tv2



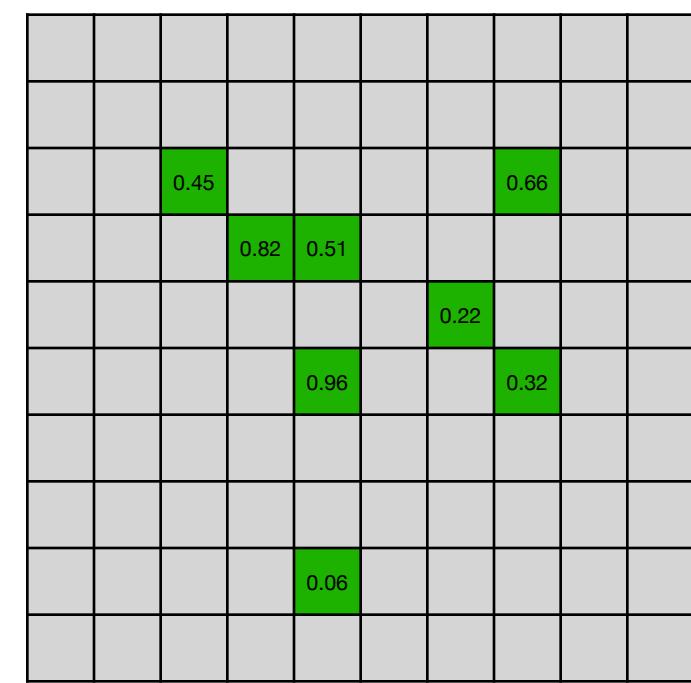
# Our Initial Hypothesis

*Having three tasks  $T_1$ ,  $T_2$  and  $T_3$*

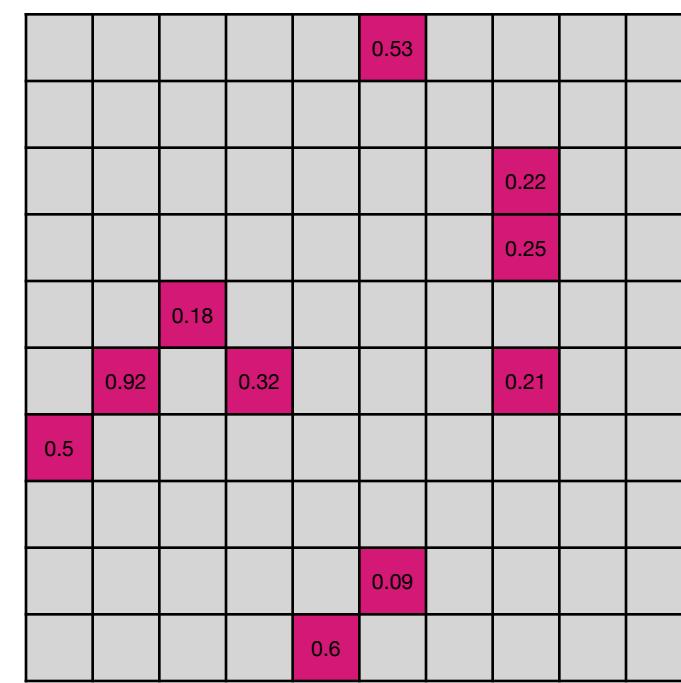
**tv1**



tv2



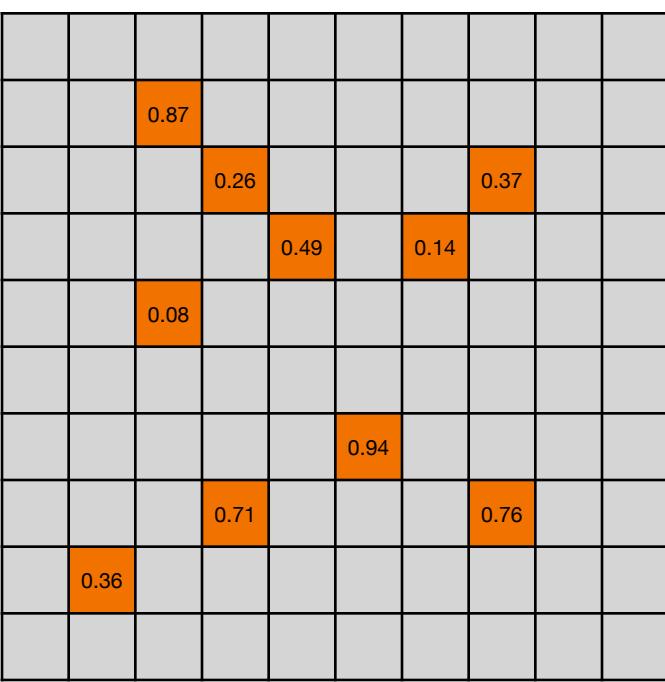
tv3



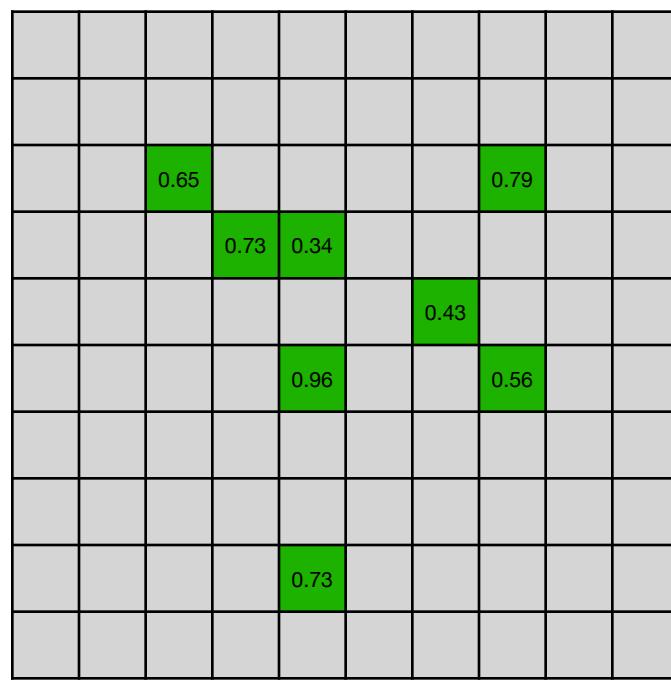
# Our Initial Hypothesis

*Having three tasks  $T_1, T_2$  and  $T_3$*

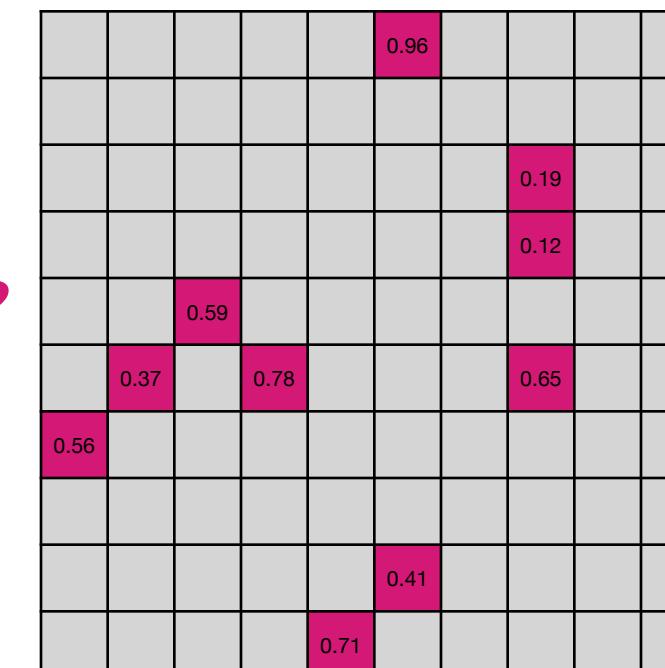
*tv1*



*tv2*



*tv3*

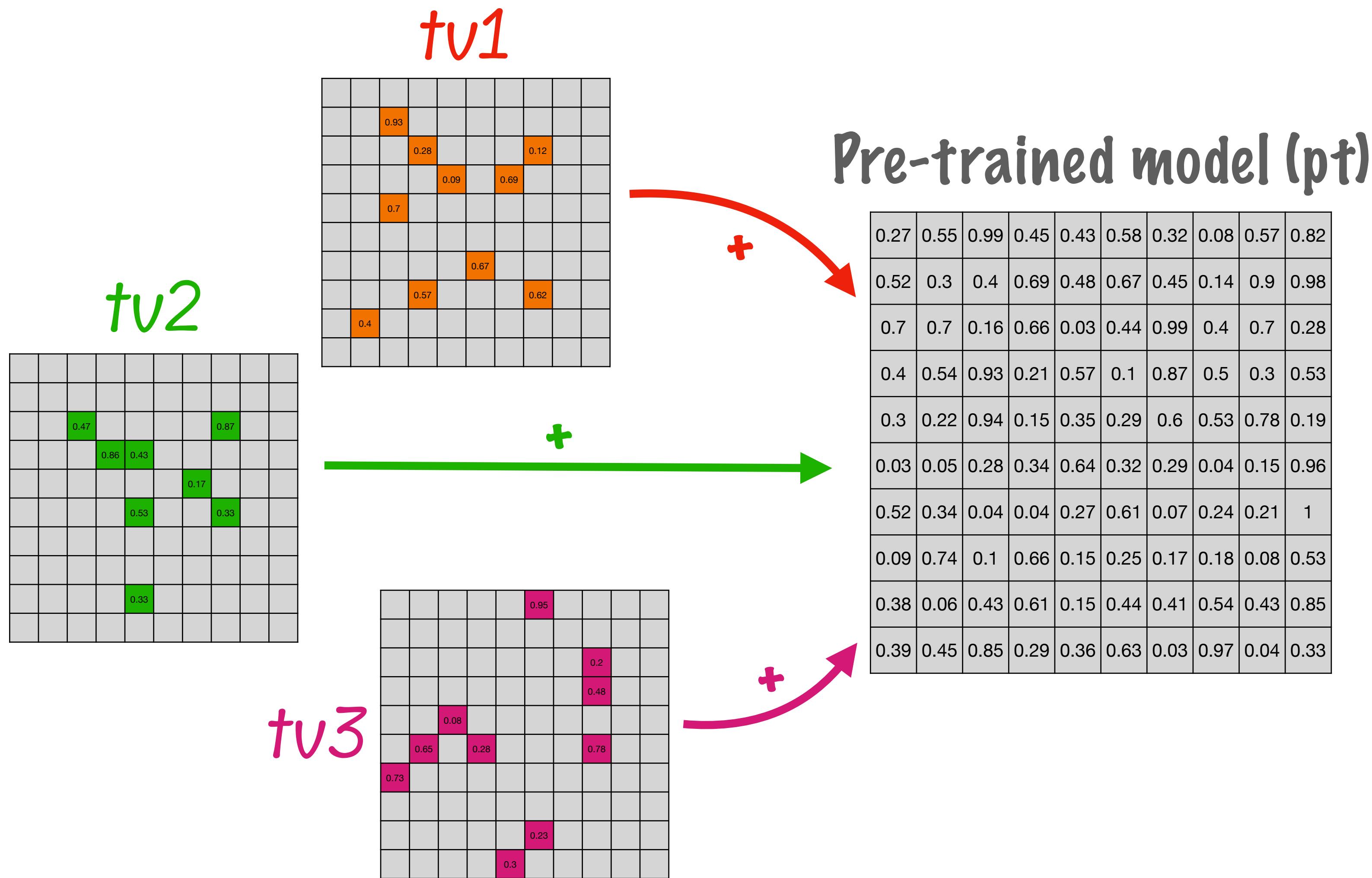


*Pre-trained model (pt)*

0.33	0.26	0.8	0.26	0.83	0.42	0.55	0.05	0.23	0.23
0.99	0.79	0.94	0.91	0.73	0.85	0.79	0.23	0.55	0.89
0.45	0.75	0.58	0.26	0.31	0.21	0.26	0.34	0.18	0.39
0.32	0.56	0.1	0.09	0.73	0.38	0.07	0.13	0.16	0.17
0.88	0.79	0.58	0.45	0.59	0.73	0.33	0.68	0.81	0.29
0.16	0.44	0.33	0.23	0.94	0.98	0.6	0.53	0.26	0.2
0.96	0.88	0.25	0.99	0.47	0.58	0.88	0.21	0.79	0.2
0.73	0.38	0.42	0.21	0.89	0.11	0.62	0.89	0.36	0.12
1	0.29	0.23	0.77	0.61	0.79	1	0.41	0.52	0.6
0.74	0.84	0.21	0.54	0.83	0.65	0.88	1	0.78	0.81

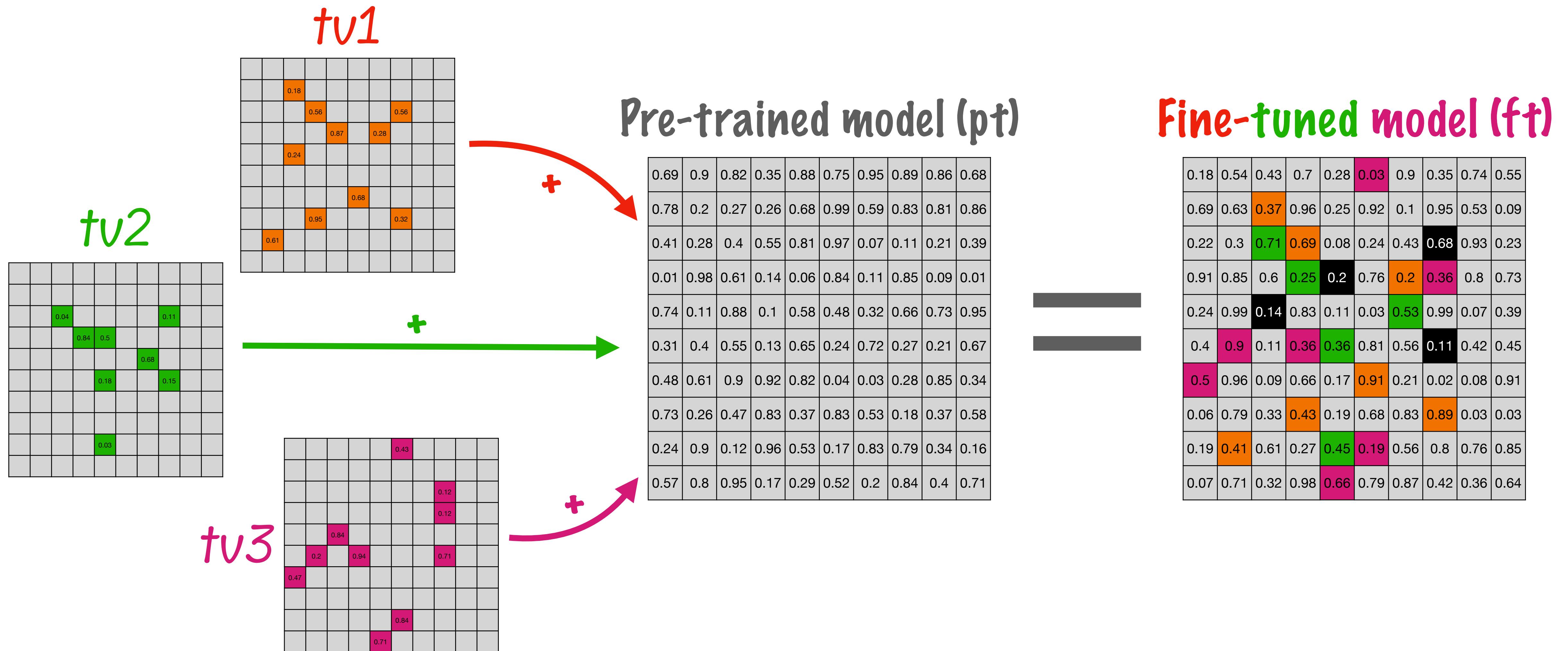
# Our Initial Hypothesis

*Having three tasks  $T_1$ ,  $T_2$  and  $T_3$*



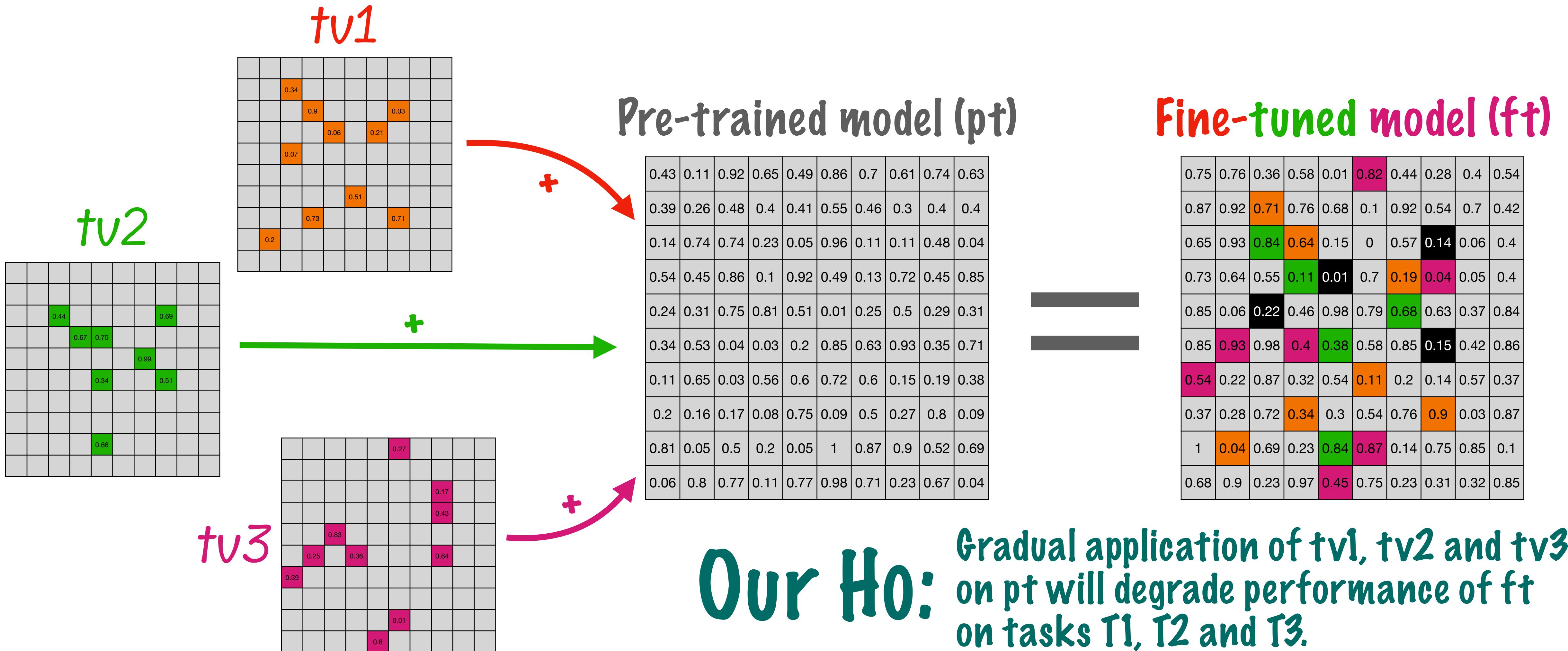
# Our Initial Hypothesis

*Having three tasks  $T_1, T_2$  and  $T_3$*



# Our Initial Hypothesis

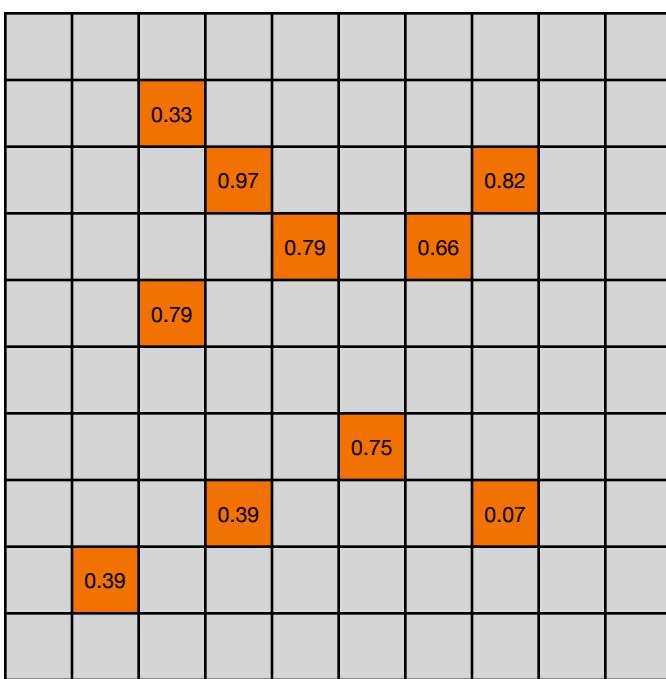
*Having three tasks  $T_1$ ,  $T_2$  and  $T_3$*



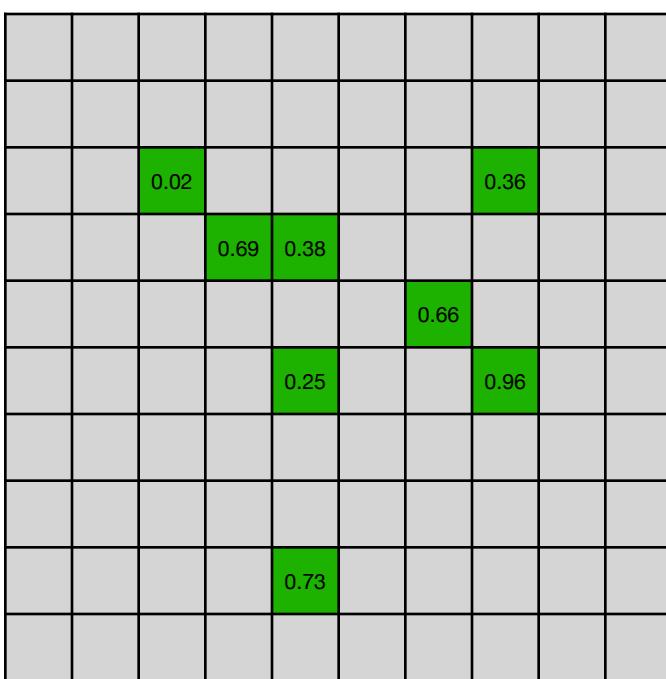
**Instead, we can do this**

# Our full Ho:

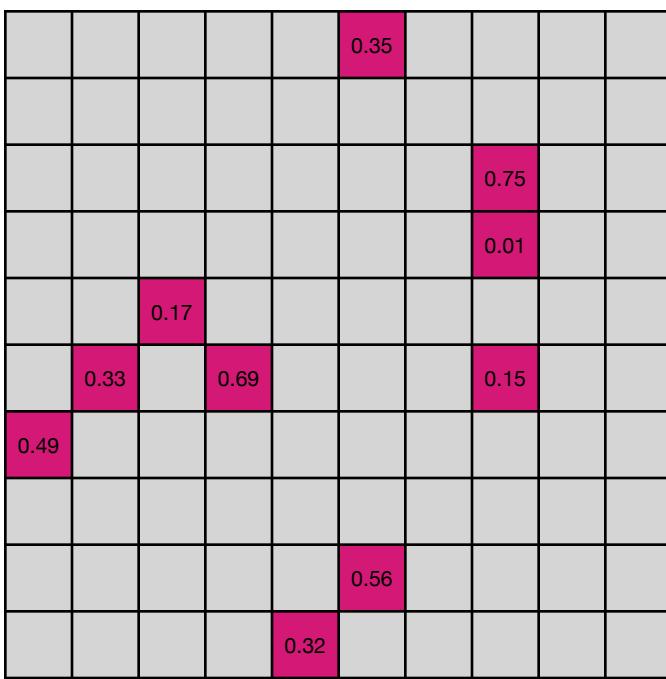
Gradual application of  $tv1$ ,  $tv2$  and  $tv3$  on  $pt$  will degrade performance of  $ft$  on tasks  $T1$ ,  $T2$  and  $T3$ . A  $ctv$  will have potential to perform well on these tasks.



tv1



tv2



tv3

# Instead, we can do this

tv1

0.34									
	0.02								
		0.75							
			0.55						
				0.86					
0.56									

## Our full Ho:

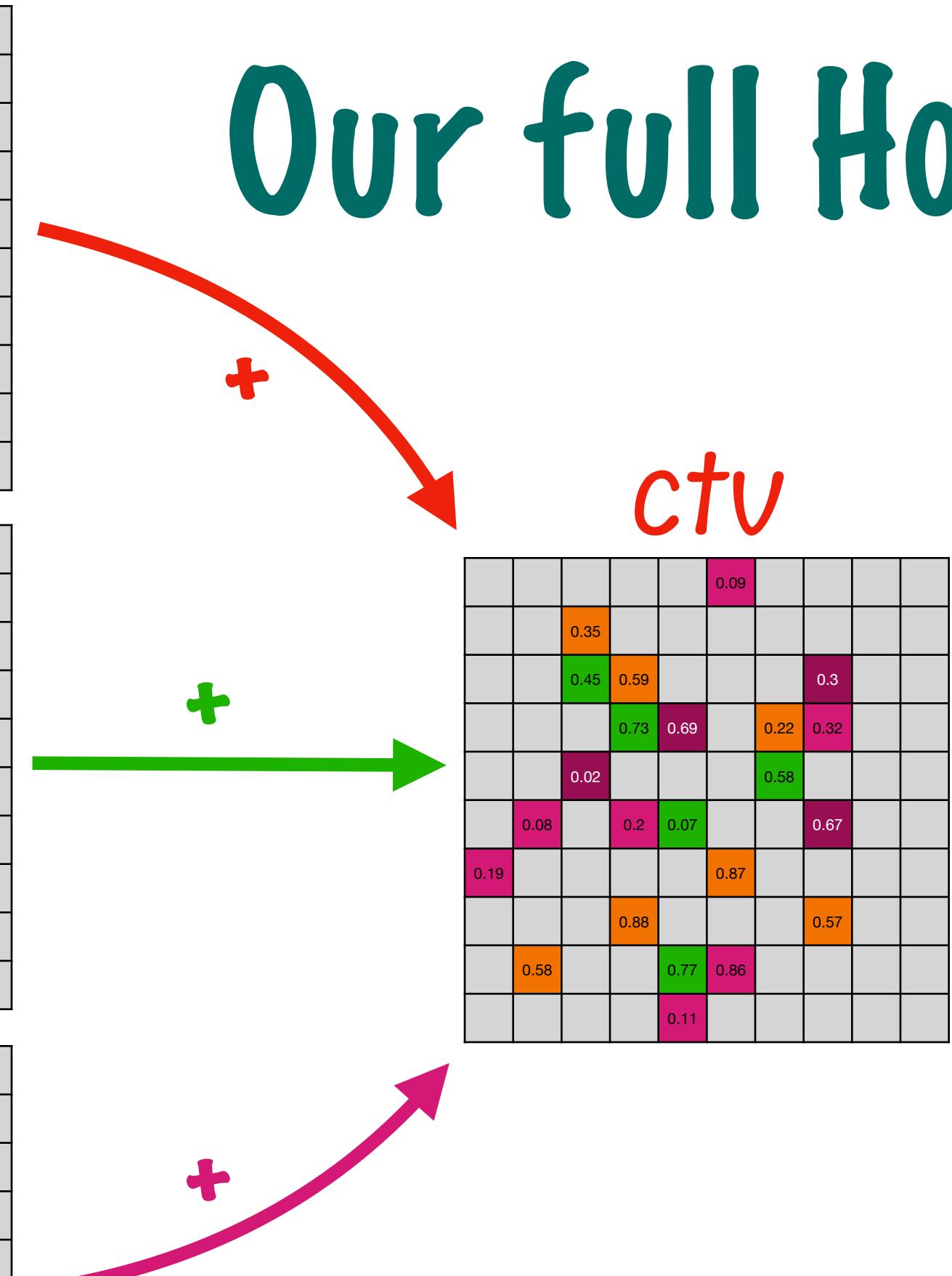
tv2

	0.07								
		0.63	0.18						
				0.65					
					0.26				
						0.69			
							0.47		

tv3

			0.14						
				0.32					
					0.81				
						0.07			
							0.73		
								0.03	
									0.14
									0.48
									0.43

Gradual application of tv1, tv2 and tv3 on pt will degrade performance of ft on tasks T1, T2 and T3. A ctv will have potential to perform well on these tasks.





# Our Initial Hypothesis

**Our full  $H_0$ :** Gradual application of  $tv1$ ,  $tv2$  and  $tv3$  on  $pt$  will degrade performance of  $ft$  on tasks  $T1$ ,  $T2$  and  $T3$ . A  $ctv$  will have potential to perform well on these tasks.

# Our Initial Hypothesis

Our full  $H_0$ :

Gradual application of  $tv1$ ,  $tv2$  and  $tv3$  on  $pt$  will degrade performance of  $ft$  on tasks  $T1$ ,  $T2$  and  $T3$ . A  $c tv$  will have potential to perform well on these tasks.

Initial results were not satisfactory so we could not prove our hypothesis.



# Our Initial Hypothesis

Our full H<sub>0</sub>:

Gradual application of tv1, tv2 and tv3 on pt will degrade performance of ft on tasks T1, T2 and T3. A ctv will have potential to perform well on these tasks.

Initial results were not satisfactory so we could not prove our hypothesis.



But during experimentation, we found something else!

# Our Initial Hypothesis

**Our full H<sub>0</sub>:**

Gradual application of tv1, tv2 and tv3 on pt will degrade performance of ft on tasks T1, T2 and T3. A ctv will have potential to perform well on these tasks.



آتے ہیں غیب سے یہ مضامیں خیال میں  
غالب سریر خامہ نوائے سروش ہے۔

Initial results were not satisfactory so we could not prove our hypothesis.



But during experimentation, we found something else!



# Sequence of Fine-tuning Models is Important

Our new  $H_0$ :

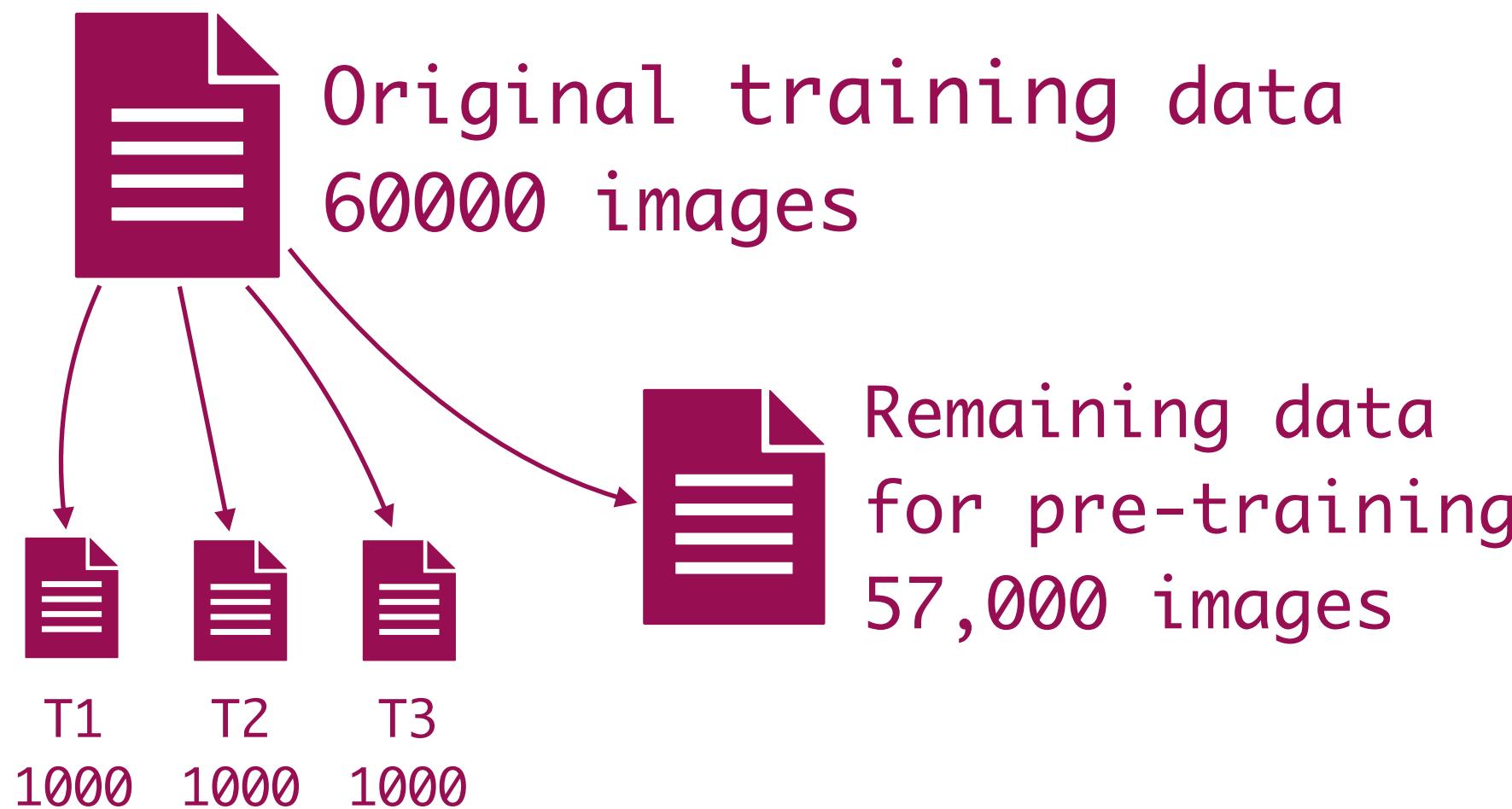
Fine-tuning a model step by step on T1, T2 and T3 yields better results as compared to fine-tuning it with combine training data of these tasks.

# Sequence of Fine-tuning Models is Important

Our new  $H_0$ :

Fine-tuning a model step by step on T1, T2 and T3 yields better results as compared to fine-tuning it with combine training data of these tasks.

We had set aside data for T1, T2 and T3 in the very beginning.



# Sequence of Fine-tuning Models is Important

Our new  $H_0$ :

Fine-tuning a model step by step on  $T_1$ ,  $T_2$  and  $T_3$  yields better results as compared to fine-tuning it with combined training data of these tasks.

We first fine-tuned with combined data for  $T_1$ ,  $T_2$  and  $T_3$ .

$$ft_{\text{comb}} = \text{train}_{T_1+T_2+T_3}(\text{pt})$$

# Sequence of Fine-tuning Models is Important

Our new  $H_0$ :

Fine-tuning a model step by step on  $T_1$ ,  $T_2$  and  $T_3$  yields better results as compared to fine-tuning it with combine training data of these tasks.

We first fine-tuned with combined data for  $T_1$ ,  $T_2$  and  $T_3$ .

$$ft_{\text{comb}} = \text{train}_{T_1+T_2+T_3}(\text{pt})$$

Then we fine-tuned with data for  $T_1$ ,  $T_2$  and  $T_3$  one after another.

$$ft_{T_1, T_2, T_3} = \text{train}_{T_3}(\text{train}_{T_2}(\text{train}_{T_1}(\text{pt})))$$

# Sequence of Fine-tuning Models is Important

Our new  $H_0$ :

Fine-tuning a model step by step on  $T_1$ ,  $T_2$  and  $T_3$  yields better results as compared to fine-tuning it with combine training data of these tasks.

We first fine-tuned with combined data for  $T_1$ ,  $T_2$  and  $T_3$ .

$$ft_{\text{comb}} = \text{train}_{T_1+T_2+T_3}(\text{pt})$$

Then we fine-tuned with data for  $T_1$ ,  $T_2$  and  $T_3$  one after another.

$$ft_{T_1, T_2, T_3} = \text{train}_{T_3}(\text{train}_{T_2}(\text{train}_{T_1}(\text{pt})))$$

Using different experiment setups we repeatedly found that

$$\text{accuracy}(ft_{T_1, T_2, T_3}) > \text{accuracy}(ft_{\text{comb}})$$

# Code and Results

# Future Plans

We may be wrong initially. We will work on our initial hypothesis by applying different techniques to compute a consolidated task vector with potential to fine-tune on multiple tasks. Techniques include:

1. Using integrated gradients to construct consolidated task vector
2. Computing attribution of affected weights in predictions to build a task vector

**\*DONE PRESENTATION\***



**ANY QUESTIONS?**

