

# SPRINT 4

## LAB 2 - DISEÑO Y AUTOMATIZACIÓN DEL PIPELINE DE ENTRENAMIENTO

Alumno: Juan Antonio Ayala Garcia  
Juananayala7@gmail.com

### 1 LANZAR 3 EXPERIMENTOS EN MLFLOW

Table Chart Evaluation Experimental											
				Metrics			Parameters				
<input type="checkbox"/>		Run Name	Created	Duration	mean_mrr10	perc_top_10	secs_per_query	model_name	normalize_emb	query_prepro_f	text_to_embed
<input type="checkbox"/>		hilarious-wolf-73	10 minutes ago	10.2min	0.476	64	0.06	hiiamsid/se...	False	modificar_...	add_gen
<input type="checkbox"/>		legendary-sheep-555	1 day ago	8.0min	0.373	52.7	0.06	stsb-xml-r-...	False	clean_quer...	add_gen
<input type="checkbox"/>		welcoming-donkey-357	1 day ago	9.3min	0.325	51.7	0.07	paraphrase...	False	clean_quer...	add_gen
<input type="checkbox"/>		unruly-cat-366	3 days ago	9.8min	0.317	47.3	0.06	stsb-xml-r-...	False	clean_quer...	get_synops...
<input type="checkbox"/>		placid-midge-610	3 days ago	5.3min	0.3	47	0.03	distiluse-ba...	False	clean_quer...	get_synops...

Captura de nuestra ventana de Mlflow

#### - MEJOR EXPERIMENTO :

- DURATION : 10.2min
- MEAN\_MRR10 : 0.476
- PERC\_TOP\_10 : 64%
- MODELO : sentence\_similarity\_spanish\_es

#### - OBSERVACIONES :

- Hemos realizado 2 Experimentos inicialmente usando dos modelos diferentes y con las funciones dadas en el ejercicio (clean\_query\_txt y get\_synopsys\_txt) en la que nuestro porcentaje fue menor de un 50%
- Realizamos 3 experimentos más probando los mismos modelos anteriores para ver si mejoramos el porcentaje, esta vez, cambiando las funciones "get\_synopsys\_txt" por "add\_gen"
- add\_gen : función usada para añadir el género de la película a la sinopsis y obviar la palabra "FILMAFFINITY", además, el resultado lo convertimos a minúsculas.