Supplementary Material for "LLMs4OM: Matching Ontologies with Large Language Models"

Hamed Babaei Giglou[0000−0003−3758−1454], Jennifer D'Souza[0000−0002−6616−9509], Felix Engel[0000−0002−3060−7052], and Sören Auer[0000−0002−0698−2864]

TIB Leibniz Information Centre for Science and Technology, Hannover, Germany
{hamed.babaei,jennifer.dsouza,felix.engel,auer}@tib.eu

Table of Contents

1 Dataset stats

We carefully chose five tracks from the OAEI campaign spanning diverse domains for our experimental configurations. The statistics for 20 datasets in five tracks

and three distinct setups—concept, concept-children, and concept-parents—are outlined in Table 1. These setups aim to determine the most effective ontology representation for OM.

Table 1. OAEI tracks and tasks statistics across source, target, and alignments.

Track	Task	Concepts		Children		Parents		Alig
		S	T	S	T	S	T	
Anatomy	Mouse-Human	2,737	3,298	482	673	1,687	3,297	1,516
Biodiv	ENVO-SWEET	6,566	4,525	2695	1,256	6,109	4,514	805
	FISH-ZOOPLANKTON	145	56	145	56	34	7	15
	ALGAE-ZOOBENTHOS	108	128	108	123	24	27	18
	TAXR-NCBI (Bacteria)	312	326	137	151	311	325	175
	TAXR-NCBI (Chromista)	2,290	2,344	933	966	2,289	2,343	1,405
	TAXR-NCBI (Fungi)	12,732	13,149	2,716	3,138	12,731	13,148	10,162
	TAXR-NCBI (Plantae)	26,302	27,013	7,324	8,003	26,301	27,012	19,914
	TAXR-NCBI (Protozoa)	501	538	147	184	500	537	357
Phenotype	DOID-ORDO	15,520	13,504	4,514	961	13,125	13,497	1,237
	HP-MP	12,786	11,928	4,387	4,439	12,646	11,498	696
CommonKG	Nell-DBpedia	134	137	0	0	0	0	129
	Yago-Wikidata	304	304	0	0	0	0	304
Bio-ML	NCIT-ORDO(disease)	15,762	8,465	2,440	2,074	7,880	8,464	4,686
	OMIM-ORDO(disease)	9,648	9,275	519	1,026	4,215	9,270	3,721
	SNOMED-FMA(body)	34,418	88,955	8,373	28,636	13,459	88,950	7,256
	SNOMED-NCIT(neoplas)	22,971	20,247	1,302	2,706	2,693	8,560	3,804
	SNOMED-NCIT(pharm)	29,500	22,136	1,300	2,284	3,527	19,030	5,803
MSE	MI-EMMO	545	903	64	232	536	704	63
	MI-MatOnto	545	825	64	114	536	793	302

2 Prompt Templates & Examples

Prompt template for C representation.

Classify if two concepts refer to the same real world entity or not (answer only yes or no).
First concept:
$\{C_s\}$
Second concept:
$\{C_t\}$
Answer:

An example:

Classify if two concepts refer to the same real world entity or not (answer only yes or no).

First concept:
cardiovascular system
Second concept:
vascular endothelium
Answer:

Prompt template for CP representation.

Classify if two concepts refer to the same real world entity or not (answer only yes or no).
First concept:
$\{C_s\}$
Parents: $\{CP\}$
Second concept:
$\{C_t\}$
Parents: $\{CP\}$
Answer:

An example:

Classify if two concepts refer to the same real world entity or not (answer only yes or no).
First concept:
cardiovascular system
Parents: organ system
Second concept:
vascular endothelium
Parents: endothelium, blood vessel tissue
Answer:

Prompt template for CC representation.

Classify if two concepts refer to the same real world entity or not (answer only yes or no).
First concept:
$\{C_s\}$
Children: $\{CC\}$
Second concept:
$\{C_t\}$
Children: $\{CC\}$
Answer:

An example:

Classify if two concepts refer to the same real world entity or not (answer only yes or no).
First concept:

cardiovascular system
Children: {*CC*}
Second concept:
vascular endothelium
Children: {*CC*}
Answer:

3 Detailed Retrieval Models Results

3.1 Anatomy Track

The detailed result of retriever models for ANATOMY track is presented in Table 2.

Table 2. Retrieval models results — ANATOMY track – Rep is the representation type.

Model	**Rep**	**Task**	$Top_k = 5$ **Results**			$Top_k = 10$ **Results**			$Top_k = 20$ **Results**		
			Prec	Rec	F1	Prec	Rec	F1	Prec	Rec	F1
Ada	C	Mouse-Human	10.6	95.71	19.09	5.4	97.49	10.23	2.73	98.55	5.31
sentence-BERT	C	Mouse-Human	10.27	92.74	18.5	5.23	94.39	9.91	2.66	95.91	5.17
SPECTER2	C	Mouse-Human	9.89	89.31	17.81	5.07	91.56	9.61	2.59	93.47	5.04
TFIDF	C	Mouse-Human	10.35	84.56	18.44	5.61	87.47	10.54	3.1	90.11	5.99
sentence-BERT	CC	Mouse-Human	9.84	88.79	17.71	5.09	91.95	9.65	2.62	94.53	5.09
SPECTER2	CC	Mouse-Human	9.48	85.62	17.08	4.94	89.25	9.37	2.56	92.41	4.98
TFIDF	CC	Mouse-Human	10.05	85.16	17.98	5.4	88.06	10.17	2.91	90.3	5.64
sentence-BERT	CP	Mouse-Human	10.01	90.37	18.03	5.18	93.47	9.81	2.65	95.58	5.15
SPECTER2	CP	Mouse-Human	9.66	87.2	17.39	5.06	91.42	9.6	2.62	94.59	5.1
TFIDF	CP	Mouse-Human	9.69	83.71	17.37	5.28	89.31	9.97	2.81	92.35	5.45

3.2 Biodiv Track

The detailed result of retriever models for BIODIV track is presented in Table 3 and Table 4.

3.3 Phenotype Track

The detailed result of retriever models for PHENOTYPE track is presented in Table 5.

3.4 CommonKG Track

The detailed result of retriever models for COMMONKG track is presented in Table 6.

3.5 Bio-ML Track

The detailed result of retriever models for Bio-ML track is presented in Table 7.

3.6 MSE Track

The detailed result of retriever models for MSE track is presented in Table 8.

4 Detailed Results of LLMs

4.1 Anatomy Track

The detailed outcomes of the LLMs4OM framework for the Anatomy track are outlined in Table 9.

4.2 Biodiv Track

The detailed outcomes of the LLMs4OM framework for the Biodiv track are outlined in Table 10, Table 11, Table 12, Table 13, Table 14, and Table 15.

4.3 Phenotype Track

The detailed outcomes of the LLMs4OM framework for the Phenotype track are outlined in Table 16, Table 17, and Table 18.

4.4 CommonKG Track

The detailed outcomes of the LLMs4OM framework for the CommonKG track are outlined in Table 19, Table 20, and Table 21.

4.5 Bio-ML Track

The detailed outcomes of the LLMs4OM framework for the bio-ml track are outlined in Table 22, Table 23, Table 24 and Table 25.

4.6 MSE Track

The detailed outcomes of the LLMs4OM framework for the mse track are outlined in Table 26, Table 27, and Table 28.

Table 3. Retrieval models results — BIODIV track — Rep is the representation type — PART 1

Model	**Rep**	**Task**	$Top_k = 5$ **Results**			$Top_k = 10$ **Results**			$Top_k = 20$ **Results**		
			Prec	Rec	F1	Prec	Rec	F1	Prec	Rec	F1
Ada	C	ENVO-SWEET	1.93	78.88	3.78	1.01	81.99	1.99	0.53	86.09	1.05
sentence-BERT	C	ENVO-SWEET	1.86	76.02	3.64	0.99	81.12	1.96	0.53	86.09	1.05
SPECTER2	C	ENVO-SWEET	1.71	69.69	3.34	0.91	74.41	1.8	0.48	78.26	0.95
TFIDF	C	ENVO-SWEET	5.52	51.93	9.98	5.33	51.93	9.67	5.06	52.05	9.23
Ada	C	FISH-ZOOPLANKTON	2.07	100.0	4.05	1.03	100.0	2.05	0.52	100.0	1.03
sentence-BERT	C	FISH-ZOOPLANKTON	1.93	93.33	3.78	0.97	93.33	1.91	0.52	100.0	1.03
SPECTER2	C	FISH-ZOOPLANKTON	2.07	100.0	4.05	1.03	100.0	2.05	0.52	100.0	1.03
TFIDF	C	FISH-ZOOPLANKTON	9.56	86.67	17.22	9.09	86.67	16.46	9.09	86.67	16.46
Ada	C	ALGAE-ZOOBENTHOS	3.15	94.44	6.09	1.67	100.0	3.28	0.83	100.0	1.65
sentence-BERT	C	ALGAE-ZOOBENTHOS	2.78	83.33	5.38	1.48	88.89	2.91	0.79	94.44	1.56
SPECTER2	C	ALGAE-ZOOBENTHOS	2.59	77.78	5.02	1.57	94.44	3.1	0.79	94.44	1.56
TFIDF	C	ALGAE-ZOOBENTHOS	16.88	72.22	27.37	15.12	72.22	25.0	13.54	72.22	22.81
Ada	C	TAXR-NCBI (Bacteria)	11.22	100.0	20.17	5.61	100.0	10.62	2.8	100.0	5.46
sentence-BERT	C	TAXR-NCBI (Bacteria)	11.22	100.0	20.17	5.61	100.0	10.62	2.8	100.0	5.46
SPECTER2	C	TAXR-NCBI (Bacteria)	11.22	100.0	20.17	5.61	100.0	10.62	2.8	100.0	5.46
TFIDF	C	TAXR-NCBI (Bacteria)	18.08	100.0	30.62	13.89	100.0	24.39	13.47	100.0	23.74
Ada	C	TAXR-NCBI (Chromista)	12.27	100.0	21.86	6.14	100.0	11.56	3.07	100.0	5.95
sentence-BERT	C	TAXR-NCBI (Chromista)	12.27	100.0	21.86	6.14	100.0	11.56	3.07	100.0	5.95
SPECTER2	C	TAXR-NCBI (Chromista)	12.24	99.79	21.81	6.14	100.0	11.56	3.07	100.0	5.95
TFIDF	C	TAXR-NCBI (Chromista)	18.32	99.29	30.93	12.52	99.93	22.26	8.77	100.0	16.13
Ada	C	TAXR-NCBI (Fungi)	15.96	100.0	27.53	7.98	100.0	14.78	3.99	100.0	7.68
sentence-BERT	C	TAXR-NCBI (Fungi)	15.96	100.0	27.53	7.98	100.0	14.78	3.99	100.0	7.68
SPECTER2	C	TAXR-NCBI (Fungi)	15.96	99.99	27.53	7.98	100.0	14.78	3.99	100.0	7.68
TFIDF	C	TAXR-NCBI (Fungi)	18.7	99.99	31.5	10.9	99.99	19.65	6.67	99.99	12.51
Ada	C	TAXR-NCBI (Plantae)	15.14	99.99	26.3	7.57	100.0	14.08	3.79	100.0	7.3
sentence-BERT	C	TAXR-NCBI (Plantae)	15.14	99.96	26.29	7.57	99.98	14.07	3.79	100.0	7.3
SPECTER2	C	TAXR-NCBI (Plantae)	15.13	99.91	26.28	7.57	99.94	14.07	3.78	99.97	7.29
TFIDF	C	TAXR-NCBI (Plantae)	17.75	99.8	30.15	10.23	99.87	18.56	6.29	99.95	11.83
Ada	C	TAXR-NCBI (Protozoa)	14.25	100.0	24.95	7.13	100.0	13.3	3.56	100.0	6.88
sentence-BERT	C	TAXR-NCBI (Protozoa)	14.25	100.0	24.95	7.13	100.0	13.3	3.56	100.0	6.88
SPECTER2	C	TAXR-NCBI (Protozoa)	14.25	100.0	24.95	7.13	100.0	13.3	3.56	100.0	6.88
TFIDF	C	TAXR-NCBI (Protozoa)	18.42	100.0	31.11	11.2	100.0	20.14	7.59	100.0	14.12

Table 4. Retrieval models results — BIODIV track — Rep is the representation type — PART 2

Model	**Rep**	**Task**	$Top_k = 5$ **Results**			$Top_k = 10$ **Results**			$Top_k = 20$ **Results**		
			Prec	Rec	F1	Prec	Rec	F1	Prec	Rec	F1
sentence-BERT	CC	ENVO-SWEET	1.46	59.63	2.85	0.81	66.21	1.6	0.45	72.92	0.89
SPECTER2	CC	ENVO-SWEET	1.29	52.8	2.53	0.73	59.38	1.44	0.4	65.59	0.8
TFIDF	CC	ENVO-SWEET	2.71	53.54	5.17	2.18	53.91	4.18	2.0	54.16	3.86
sentence-BERT	CC	FISH-ZOOPLANKTON	2.07	100.0	4.05	1.03	100.0	2.05	0.52	100.0	1.03
SPECTER2	CC	FISH-ZOOPLANKTON	1.66	80.0	3.24	0.9	86.67	1.77	0.45	86.67	0.89
TFIDF	CC	FISH-ZOOPLANKTON	2.71	80.0	5.24	1.53	86.67	3.0	0.88	93.33	1.75
sentence-BERT	CC	ALGAE-ZOOBENTHOS	2.96	88.89	5.73	1.57	94.44	3.1	0.83	100.0	1.65
SPECTER2	CC	ALGAE-ZOOBENTHOS	2.41	72.22	4.66	1.3	77.78	2.55	0.74	88.89	1.47
TFIDF	CC	ALGAE-ZOOBENTHOS	4.3	72.22	8.12	2.53	83.33	4.92	1.82	94.44	3.58
sentence-BERT	CC	TAXR-NCBI (Bacteria)	11.22	100.0	20.17	5.61	100.0	10.62	2.8	100.0	5.46
SPECTER2	CC	TAXR-NCBI (Bacteria)	11.22	100.0	20.17	5.61	100.0	10.62	2.8	100.0	5.46
TFIDF	CC	TAXR-NCBI (Bacteria)	13.06	100.0	23.1	9.68	100.0	17.65	9.0	100.0	16.52
sentence-BERT	CC	TAXR-NCBI (Chromista)	12.27	100.0	21.86	6.14	100.0	11.56	3.07	100.0	5.95
SPECTER2	CC	TAXR-NCBI (Chromista)	12.25	99.86	21.83	6.14	100.0	11.56	3.07	100.0	5.95
TFIDF	CC	TAXR-NCBI (Chromista)	13.86	99.15	24.32	9.23	99.93	16.91	6.63	100.0	12.43
sentence-BERT	CC	TAXR-NCBI (Fungi)	15.96	100.0	27.53	7.98	100.0	14.78	3.99	100.0	7.68
SPECTER2	CC	TAXR-NCBI (Fungi)	15.96	99.99	27.53	7.98	100.0	14.78	3.99	100.0	7.68
TFIDF	CC	TAXR-NCBI (Fungi)	16.73	99.99	28.66	9.41	99.99	17.19	5.69	99.99	10.77
sentence-BERT	CC	TAXR-NCBI (Plantae)	15.14	99.96	26.29	7.57	99.98	14.07	3.79	99.99	7.29
SPECTER2	CC	TAXR-NCBI (Plantae)	15.13	99.89	26.27	7.57	99.92	14.07	3.78	99.96	7.29
TFIDF	CC	TAXR-NCBI (Plantae)	15.58	99.8	26.96	8.46	99.87	15.59	4.89	99.95	9.32
sentence-BERT	CC	TAXR-NCBI (Protozoa)	14.25	100.0	24.95	7.13	100.0	13.3	3.56	100.0	6.88
SPECTER2	CC	TAXR-NCBI (Protozoa)	14.25	100.0	24.95	7.13	100.0	13.3	3.56	100.0	6.88
TFIDF	CC	TAXR-NCBI (Protozoa)	15.8	100.0	27.29	9.51	100.0	17.38	6.46	100.0	12.13
sentence-BERT	CP	ENVO-SWEET	1.49	60.87	2.91	0.86	69.94	1.69	0.48	78.39	0.96
SPECTER2	CP	ENVO-SWEET	1.33	54.04	2.59	0.76	61.99	1.5	0.42	68.32	0.83
TFIDF	CP	ENVO-SWEET	2.18	53.79	4.2	1.45	58.26	2.84	1.1	61.61	2.17
sentence-BERT	CP	FISH-ZOOPLANKTON	1.93	93.33	3.78	0.97	93.33	1.91	0.52	100.0	1.03
SPECTER2	CP	FISH-ZOOPLANKTON	1.66	80.0	3.24	0.97	93.33	1.91	0.52	100.0	1.03
TFIDF	CP	FISH-ZOOPLANKTON	6.67	73.33	12.22	5.41	80.0	10.13	5.73	86.67	10.74
sentence-BERT	CP	ALGAE-ZOOBENTHOS	2.78	83.33	5.38	1.48	88.89	2.91	0.83	100.0	1.65
SPECTER2	CP	ALGAE-ZOOBENTHOS	2.78	83.33	5.38	1.48	88.89	2.91	0.83	100.0	1.65
TFIDF	CP	ALGAE-ZOOBENTHOS	12.37	66.67	20.87	9.92	72.22	17.45	8.55	72.22	15.29
sentence-BERT	CP	TAXR-NCBI (Bacteria)	11.22	100.0	20.17	5.61	100.0	10.62	2.8	100.0	5.46
SPECTER2	CP	TAXR-NCBI (Bacteria)	11.22	100.0	20.17	5.61	100.0	10.62	2.8	100.0	5.46
TFIDF	CP	TAXR-NCBI (Bacteria)	14.06	100.0	24.65	10.25	100.0	18.59	9.72	100.0	17.71
sentence-BERT	CP	TAXR-NCBI (Chromista)	12.24	99.72	21.8	6.14	100.0	11.56	3.07	100.0	5.95
SPECTER2	CP	TAXR-NCBI (Chromista)	12.19	99.36	21.72	6.11	99.64	11.52	3.06	99.72	5.94
TFIDF	CP	TAXR-NCBI (Chromista)	14.58	99.57	25.43	9.61	100.0	17.53	6.8	100.0	12.73
sentence-BERT	CP	TAXR-NCBI (Fungi)	15.94	99.88	27.5	7.98	99.93	14.77	3.99	100.0	7.68
SPECTER2	CP	TAXR-NCBI (Fungi)	15.88	99.49	27.39	7.95	99.6	14.72	3.98	99.69	7.65
TFIDF	CP	TAXR-NCBI (Fungi)	16.95	99.97	28.99	9.53	99.99	17.4	5.76	99.99	10.9
sentence-BERT	CP	TAXR-NCBI (Plantae)	15.12	99.82	26.26	7.56	99.9	14.06	3.78	99.96	7.29
SPECTER2	CP	TAXR-NCBI (Plantae)	15.01	99.1	26.06	7.52	99.35	13.98	3.77	99.52	7.26
TFIDF	CP	TAXR-NCBI (Plantae)	15.93	99.83	27.48	8.93	99.9	16.4	5.55	99.95	10.52
sentence-BERT	CP	TAXR-NCBI (Protozoa)	14.25	100.0	24.95	7.13	100.0	13.3	3.56	100.0	6.88
SPECTER2	CP	TAXR-NCBI (Protozoa)	14.25	100.0	24.95	7.13	100.0	13.3	3.56	100.0	6.88
TFIDF	CP	TAXR-NCBI (Protozoa)	16.8	100.0	28.77	10.33	100.0	18.73	7.13	100.0	13.31

Table 5. Retrieval models results — PHENOTYPE track – Rep is the representation type.

Model	**Rep**	**Task**	$Top_k = 5$ **Results**			$Top_k = 10$ **Results**			$Top_k = 20$ **Results**		
			Prec	Rec	F1	Prec	Rec	F1	Prec	Rec	F1
Ada	C	DOID-ORDO	1.61	100.0	3.17	0.8	100.0	1.6	0.4	100.0	0.8
sentence-BERT	C	DOID-ORDO	1.61	100.0	3.17	0.8	100.0	1.6	0.4	100.0	0.8
SPECTER2	C	DOID-ORDO	1.61	99.92	3.17	0.8	100.0	1.6	0.4	100.0	0.8
TFIDF	C	DOID-ORDO	2.01	98.54	3.93	1.06	99.6	2.09	0.56	99.68	1.12
Ada	C	HP-MP	1.09	99.71	2.15	0.54	99.71	1.08	0.27	99.86	0.54
sentence-BERT	C	HP-MP	1.09	99.71	2.15	0.54	99.71	1.08	0.27	99.71	0.54
SPECTER2	C	HP-MP	1.09	99.86	2.15	0.54	99.86	1.08	0.27	99.86	0.54
TFIDF	C	HP-MP	1.31	99.43	2.58	0.69	99.43	1.36	0.37	99.57	0.73
sentence-BERT	CC	DOID-ORDO	1.45	90.14	2.86	0.74	92.08	1.47	0.38	93.94	0.75
SPECTER2	CC	DOID-ORDO	1.54	95.55	3.03	0.79	97.66	1.56	0.4	98.87	0.79
TFIDF	CC	DOID-ORDO	1.86	97.17	3.65	0.97	98.46	1.93	0.51	99.19	1.02
sentence-BERT	CC	HP-MP	0.94	86.49	1.86	0.48	88.07	0.95	0.25	90.66	0.49
SPECTER2	CC	HP-MP	1.0	91.67	1.97	0.51	94.54	1.02	0.26	96.41	0.52
TFIDF	CC	HP-MP	1.19	96.98	2.36	0.62	98.13	1.24	0.33	98.99	0.66
sentence-BERT	CP	DOID-ORDO	1.51	93.61	2.97	0.77	96.12	1.53	0.39	98.14	0.79
SPECTER2	CP	DOID-ORDO	1.48	91.84	2.91	0.77	95.31	1.52	0.39	97.33	0.78
TFIDF	CP	DOID-ORDO	1.69	92.56	3.32	0.9	96.28	1.78	0.47	98.14	0.94
sentence-BERT	CP	HP-MP	1.02	93.25	2.01	0.53	96.55	1.05	0.27	98.13	0.53
SPECTER2	CP	HP-MP	1.04	95.26	2.05	0.53	97.41	1.05	0.27	99.28	0.54
TFIDF	CP	HP-MP	1.04	91.95	2.06	0.56	96.55	1.11	0.29	98.42	0.58

Table 6. Retrieval models results — CommonKG track – Rep is the representation type.

Model	**Rep**	**Task**	$Top_k = 5$ **Results**			$Top_k = 10$ **Results**			$Top_k = 20$ **Results**		
			Prec	Rec	F1	Prec	Rec	F1	Prec	Rec	F1
Ada	C	Nell-DBpedia	18.96	98.45	31.79	9.55	99.22	17.43	4.81	100.0	9.18
sentence-BERT	C	Nell-DBpedia	18.96	98.45	31.79	9.55	99.22	17.43	4.81	100.0	9.18
SPECTER2	C	Nell-DBpedia	18.51	96.12	31.04	9.25	96.12	16.88	4.63	96.12	8.83
TFIDF	C	Nell-DBpedia	100.0	78.29	87.83	100.0	78.29	87.83	100.0	78.29	87.83
Ada	C	Yago-Wikidata	19.67	98.36	32.79	9.87	98.68	17.94	4.95	99.01	9.43
sentence-BERT	C	Yago-Wikidata	19.47	97.37	32.46	9.84	98.36	17.88	4.97	99.34	9.46
SPECTER2	C	Yago-Wikidata	17.76	88.82	29.61	9.14	91.45	16.63	4.69	93.75	8.93
TFIDF	C	Yago-Wikidata	73.94	40.13	52.03	72.62	40.13	51.69	72.62	40.13	51.69
sentence-BERT	CC	Nell-DBpedia	18.96	98.45	31.79	9.55	99.22	17.43	4.81	100.0	9.18
SPECTER2	CC	Nell-DBpedia	18.51	96.12	31.04	9.25	96.12	16.88	4.63	96.12	8.83
TFIDF	CC	Nell-DBpedia	100.0	78.29	87.83	100.0	78.29	87.83	100.0	78.29	87.83
sentence-BERT	CC	Yago-Wikidata	19.47	97.37	32.46	9.84	98.36	17.88	4.97	99.34	9.46
SPECTER2	CC	Yago-Wikidata	17.76	88.82	29.61	9.14	91.45	16.63	4.69	93.75	8.93
TFIDF	CC	Yago-Wikidata	73.94	40.13	52.03	72.62	40.13	51.69	72.62	40.13	51.69
sentence-BERT	CP	Nell-DBpedia	18.96	98.45	31.79	9.55	99.22	17.43	4.81	100.0	9.18
SPECTER2	CP	Nell-DBpedia	18.51	96.12	31.04	9.25	96.12	16.88	4.63	96.12	8.83
TFIDF	CP	Nell-DBpedia	100.0	78.29	87.83	100.0	78.29	87.83	100.0	78.29	87.83
sentence-BERT	CP	Yago-Wikidata	19.47	97.37	32.46	9.84	98.36	17.88	4.97	99.34	9.46
SPECTER2	CP	Yago-Wikidata	17.76	88.82	29.61	9.14	91.45	16.63	4.69	93.75	8.93
TFIDF	CP	Yago-Wikidata	73.94	40.13	52.03	72.62	40.13	51.69	72.62	40.13	51.69

Table 7. Retrieval models results — Bio-ML track – Rep is the representation type.

Model	**Rep**	**Task**	$Top_k = 5$ **Results**			$Top_k = 10$ **Results**			$Top_k = 20$ **Results**		
			Prec	Rec	F1	Prec	Rec	F1	Prec	Rec	F1
Ada	C	NCIT-DOID(disease)	5.47	91.91	10.32	2.8	94.3	5.45	1.43	95.92	2.81
sentence-BERT	C	NCIT-DOID(disease)	5.37	90.25	10.13	2.74	92.23	5.33	1.4	93.85	2.75
SPECTER2	C	NCIT-DOID(disease)	5.4	90.74	10.18	2.76	92.83	5.36	1.41	94.6	2.77
TFIDF	C	NCIT-DOID(disease)	5.41	81.2	10.15	2.95	84.96	5.69	1.61	87.73	3.16
Ada	C	OMIM-ORDO(disease)	5.62	72.83	10.43	2.93	76.05	5.65	1.52	78.66	2.98
sentence-BERT	C	OMIM-ORDO(disease)	5.51	71.49	10.24	2.89	74.95	5.57	1.49	77.32	2.93
SPECTER2	C	OMIM-ORDO(disease)	5.47	70.89	10.15	2.84	73.72	5.47	1.48	76.7	2.9
TFIDF	C	OMIM-ORDO(disease)	5.56	69.44	10.3	2.97	73.07	5.71	1.58	76.08	3.09
Ada	C	SNOMED-FMA(body)	3.4	80.54	6.52	1.81	85.72	3.54	0.94	89.14	1.86
sentence-BERT	C	SNOMED-FMA(body)	3.15	74.66	6.04	1.67	79.37	3.28	0.87	82.87	1.73
SPECTER2	C	SNOMED-FMA(body)	2.22	52.63	4.26	1.29	60.98	2.52	0.72	68.61	1.43
TFIDF	C	SNOMED-FMA(body)	1.3	30.29	2.49	0.8	37.31	1.57	0.54	50.32	1.07
Ada	C	SNOMED-NCIT(neoplas)	2.74	82.62	5.3	1.43	86.65	2.82	0.74	89.04	1.46
sentence-BERT	C	SNOMED-NCIT(neoplas)	2.64	79.65	5.11	1.38	83.39	2.72	0.71	86.2	1.42
SPECTER2	C	SNOMED-NCIT(neoplas)	2.6	78.55	5.04	1.37	82.65	2.69	0.72	86.38	1.42
TFIDF	C	SNOMED-NCIT(neoplas)	2.21	66.11	4.27	1.18	70.27	2.32	0.63	74.37	1.24
Ada	C	SNOMED-NCIT(pharm)	3.75	95.31	7.22	1.89	96.16	3.71	0.95	96.95	1.89
sentence-BERT	C	SNOMED-NCIT(pharm)	3.65	92.87	7.03	1.84	93.73	3.62	0.93	94.55	1.84
SPECTER2	C	SNOMED-NCIT(pharm)	3.52	89.49	6.77	1.79	91.09	3.51	0.91	92.28	1.8
TFIDF	C	SNOMED-NCIT(pharm)	2.89	73.39	5.57	1.82	88.78	3.57	0.98	89.68	1.94
sentence-BERT	CC	NCIT-DOID(disease)	4.67	78.57	8.82	2.45	82.52	4.77	1.27	85.64	2.51
SPECTER2	CC	NCIT-DOID(disease)	4.81	80.86	9.08	2.56	86.06	4.97	1.34	90.25	2.64
TFIDF	CC	NCIT-DOID(disease)	5.17	81.09	9.72	2.82	85.81	5.47	1.52	88.8	2.98
sentence-BERT	CC	OMIM-ORDO(disease)	5.18	67.21	9.63	2.74	71.08	5.28	1.43	74.36	2.81
SPECTER2	CC	OMIM-ORDO(disease)	5.35	69.34	9.93	2.8	72.61	5.39	1.46	75.89	2.87
TFIDF	CC	OMIM-ORDO(disease)	5.5	69.5	10.19	2.93	73.21	5.64	1.55	76.05	3.04
sentence-BERT	CC	SNOMED-FMA(body)	1.99	47.23	3.82	1.16	55.03	2.27	0.67	63.27	1.32
SPECTER2	CC	SNOMED-FMA(body)	1.58	37.49	3.03	0.95	45.29	1.87	0.56	52.99	1.11
TFIDF	CC	SNOMED-FMA(body)	1.62	38.01	3.11	0.99	46.4	1.94	0.58	54.22	1.15
sentence-BERT	CC	SNOMED-NCIT(neoplas)	2.49	75.32	4.83	1.31	79.1	2.58	0.68	82.44	1.35
SPECTER2	CC	SNOMED-NCIT(neoplas)	2.51	75.87	4.86	1.33	80.18	2.61	0.7	84.07	1.38
TFIDF	CC	SNOMED-NCIT(neoplas)	2.2	65.9	4.25	1.17	70.35	2.31	0.63	74.74	1.24
sentence-BERT	CC	SNOMED-NCIT(pharm)	3.57	90.75	6.87	1.81	91.87	3.54	0.91	92.97	1.81
SPECTER2	CC	SNOMED-NCIT(pharm)	3.39	86.15	6.52	1.73	87.75	3.39	0.87	88.92	1.73
TFIDF	CC	SNOMED-NCIT(pharm)	3.37	85.58	6.49	1.72	87.11	3.37	0.92	88.13	1.82
sentence-BERT	CP	NCIT-DOID(disease)	5.1	85.81	9.63	2.67	89.82	5.19	1.38	92.74	2.72
SPECTER2	CP	NCIT-DOID(disease)	5.21	87.58	9.83	2.72	91.36	5.28	1.4	94.52	2.77
TFIDF	CP	NCIT-DOID(disease)	4.91	79.66	9.25	2.67	85.0	5.17	1.45	89.46	2.85
sentence-BERT	CP	OMIM-ORDO(disease)	5.43	70.44	10.09	2.86	74.07	5.5	1.49	77.05	2.92
SPECTER2	CP	OMIM-ORDO(disease)	5.28	68.42	9.8	2.81	72.83	5.41	1.47	76.24	2.88
TFIDF	CP	OMIM-ORDO(disease)	5.53	69.12	10.24	2.97	73.15	5.7	1.58	76.48	3.09
sentence-BERT	CP	SNOMED-FMA(body)	2.68	63.45	5.13	1.49	70.53	2.91	0.81	76.67	1.6
SPECTER2	CP	SNOMED-FMA(body)	2.01	47.7	3.86	1.19	56.35	2.33	0.69	65.09	1.36
TFIDF	CP	SNOMED-FMA(body)	1.06	24.74	2.03	0.75	34.9	1.46	0.55	50.91	1.08
sentence-BERT	CP	SNOMED-NCIT(neoplas)	2.46	74.37	4.77	1.31	79.07	2.58	0.68	82.57	1.36
SPECTER2	CP	SNOMED-NCIT(neoplas)	2.46	74.19	4.76	1.31	79.02	2.57	0.69	83.02	1.36
TFIDF	CP	SNOMED-NCIT(neoplas)	2.18	65.43	4.22	1.17	70.16	2.31	0.62	74.13	1.23
sentence-BERT	CP	SNOMED-NCIT(pharm)	3.02	76.67	5.8	1.6	81.58	3.15	0.84	85.47	1.66
SPECTER2	CP	SNOMED-NCIT(pharm)	2.87	72.98	5.53	1.53	77.79	3.0	0.8	81.46	1.59
TFIDF	CP	SNOMED-NCIT(pharm)	2.24	56.9	4.31	1.42	72.2	2.79	0.81	82.1	1.6

Table 8. Retrieval models results — MSE track – Rep is the representation type.

Model	**Rep**	**Task**	$Top_k = 5$ **Results**			$Top_k = 10$ **Results**			$Top_k = 20$ **Results**		
			Prec	Rec	F1	Prec	Rec	F1	Prec	Rec	F1
Ada	C	MI-EMMO	10.85	95.24	19.48	10.95	95.24	19.64	10.91	95.24	19.58
sentence-BERT	C	MI-EMMO	11.56	100.0	20.72	11.56	100.0	20.72	11.56	100.0	20.72
SPECTER2	C	MI-EMMO	11.19	96.83	20.07	11.19	96.83	20.07	11.19	96.83	20.07
TFIDF	C	MI-EMMO	22.92	87.3	36.3	22.92	87.3	36.3	22.92	87.3	36.3
Ada	C	MI-MatOnto	3.74	33.77	6.74	2.22	40.07	4.21	1.39	50.0	2.7
sentence-BERT	C	MI-MatOnto	5.43	49.01	9.78	3.28	59.27	6.22	1.77	63.91	3.45
SPECTER2	C	MI-MatOnto	2.9	26.16	5.22	1.45	26.16	2.75	0.78	28.15	1.52
TFIDF	C	MI-MatOnto	5.85	22.19	9.25	3.8	22.19	6.48	2.87	22.19	5.08
sentence-BERT	CC	MI-EMMO	7.89	68.25	14.14	7.89	68.25	14.14	7.89	68.25	14.14
SPECTER2	CC	MI-EMMO	6.61	57.14	11.84	6.61	57.14	11.84	6.61	57.14	11.84
TFIDF	CC	MI-EMMO	15.69	63.49	25.16	15.69	63.49	25.16	15.69	63.49	25.16
sentence-BERT	CC	MI-MatOnto	5.25	47.35	9.45	3.28	59.27	6.22	1.8	64.9	3.5
SPECTER2	CC	MI-MatOnto	2.46	22.19	4.43	1.34	24.17	2.54	0.77	27.81	1.5
TFIDF	CC	MI-MatOnto	4.75	22.19	7.82	2.89	22.19	5.11	1.97	22.19	3.61
sentence-BERT	CP	MI-EMMO	5.89	53.97	10.62	6.05	53.97	10.88	5.96	53.97	10.74
SPECTER2	CP	MI-EMMO	3.96	34.92	7.12	4.0	34.92	7.18	3.98	34.92	7.14
TFIDF	CP	MI-EMMO	4.66	41.27	8.37	4.66	41.27	8.37	4.66	41.27	8.37
sentence-BERT	CP	MI-MatOnto	5.28	47.68	9.51	3.03	54.64	5.74	1.65	59.6	3.21
SPECTER2	CP	MI-MatOnto	5.32	48.01	9.58	3.05	54.97	5.77	1.67	60.26	3.25
TFIDF	CP	MI-MatOnto	2.91	23.18	5.17	1.71	25.17	3.21	1.1	27.48	2.11

Table 9. LLM models results — ANATOMY track – Rep is the representation type. Retriever model Top-k is set to 5.

Model	Rep	Task	Results		
			Prec	Rec	F1
GPT-3.5 + Ada	C	Mouse-Human	90.82	87.47	89.11
Falcon-7B + Ada	C	Mouse-Human	87.71	87.07	87.39
Falcon-7B + BERT	C	Mouse-Human	97.9	73.75	84.12
LLaMA-2-7B + Ada	C	Mouse-Human	87.71	87.07	87.39
LLaMA-2-7B + BERT	C	Mouse-Human	97.9	73.75	84.12
MPT-7B + Ada	C	Mouse-Human	87.71	87.07	87.39
MPT-7B + BERT	C	Mouse-Human	97.9	73.75	84.12
Mamba-2.8B + Ada	C	Mouse-Human	79.69	73.48	76.46
Mamba-2.8B + BERT	C	Mouse-Human	96.01	61.94	75.3
Mistral-7B + Ada	C	Mouse-Human	91.32	86.74	88.97
Mistral-7B + BERT	C	Mouse-Human	98.49	73.22	84.0
Vicuna-7B + Ada	C	Mouse-Human	87.09	85.03	86.05
Vicuna-7B + BERT	C	Mouse-Human	97.48	71.57	82.54
GPT-3.5 + Ada	CC	Mouse-Human	90.83	86.87	88.81
Falcon-7B + Ada	CC	Mouse-Human	87.55	86.74	87.14
Falcon-7B + BERT	CC	Mouse-Human	97.89	73.42	83.91
LLaMA-2-7B + Ada	CC	Mouse-Human	87.51	86.87	87.19
LLaMA-2-7B + BERT	CC	Mouse-Human	97.89	73.55	83.99
MPT-7B + Ada	CC	Mouse-Human	87.7	87.01	87.35
MPT-7B + BERT	CC	Mouse-Human	97.89	73.61	84.04
Mamba-2.8B + Ada	CC	Mouse-Human	80.53	74.21	77.24
Mamba-2.8B + BERT	CC	Mouse-Human	96.68	63.46	76.62
Mistral-7B + Ada	CC	Mouse-Human	95.11	76.91	85.05
Mistral-7B + BERT	CC	Mouse-Human	98.93	67.08	79.95
Vicuna-7B + Ada	CC	Mouse-Human	86.46	84.7	85.57
Vicuna-7B + BERT	CC	Mouse-Human	97.83	71.24	82.44
GPT-3.5 + Ada	CP	Mouse-Human	91.89	84.5	88.04
Falcon-7B + Ada	CP	Mouse-Human	87.71	87.07	87.39
Falcon-7B + BERT	CP	Mouse-Human	97.9	73.75	84.12
LLaMA-2-7B + Ada	CP	Mouse-Human	87.71	87.07	87.39
LLaMA-2-7B + BERT	CP	Mouse-Human	97.9	73.75	84.12
MPT-7B + Ada	CP	Mouse-Human	87.71	87.07	87.39
MPT-7B + BERT	CP	Mouse-Human	97.9	73.75	84.12
Mamba-2.8B + Ada	CP	Mouse-Human	80.27	75.13	77.61
Mamba-2.8B + BERT	CP	Mouse-Human	96.88	63.39	76.63
Mistral-7B + Ada	CP	Mouse-Human	92.2	81.07	86.28
Mistral-7B + BERT	CP	Mouse-Human	98.97	69.92	81.95
Vicuna-7B + Ada	CP	Mouse-Human	85.74	83.31	84.51
Vicuna-7B + BERT	CP	Mouse-Human	97.51	69.72	81.31

Table 10. LLM models results — BIODIV track – Rep is the representation type. Retriever model Top-k is set to 5. PART 1

Model	Rep	Task	Results		
			Prec	Rec	F1
GPT-3.5 + Ada	C	ENVO-SWEET	59.62	46.21	52.06
Falcon-7B + Ada	C	ENVO-SWEET	55.02	53.79	54.4
Falcon-7B + BERT	C	ENVO-SWEET	81.63	29.81	43.68
LLaMA-2-7B + Ada	C	ENVO-SWEET	55.02	53.79	54.4
LLaMA-2-7B + BERT	C	ENVO-SWEET	81.63	29.81	43.68
MPT-7B + Ada	C	ENVO-SWEET	55.02	53.79	54.4
MPT-7B + BERT	C	ENVO-SWEET	81.63	29.81	43.68
Mamba-2.8B + Ada	C	ENVO-SWEET	50.61	46.21	48.31
Mamba-2.8B + BERT	C	ENVO-SWEET	81.25	25.84	39.21
Mistral-7B + Ada	C	ENVO-SWEET	59.01	51.68	55.1
Mistral-7B + BERT	C	ENVO-SWEET	82.76	29.81	43.84
Vicuna-7B + Ada	C	ENVO-SWEET	51.61	47.83	49.65
Vicuna-7B + BERT	C	ENVO-SWEET	81.44	26.71	40.22
GPT-3.5 + Ada	C	FISH-ZOOPLANKTON	100.0	73.33	84.62
Falcon-7B + Ada	C	FISH-ZOOPLANKTON	100.0	80.0	88.89
Falcon-7B + BERT	C	FISH-ZOOPLANKTON	100.0	53.33	69.57
LLaMA-2-7B + Ada	C	FISH-ZOOPLANKTON	100.0	80.0	88.89
LLaMA-2-7B + BERT	C	FISH-ZOOPLANKTON	100.0	53.33	69.57
MPT-7B + Ada	C	FISH-ZOOPLANKTON	100.0	80.0	88.89
MPT-7B + BERT	C	FISH-ZOOPLANKTON	100.0	53.33	69.57
Mamba-2.8B + Ada	C	FISH-ZOOPLANKTON	90.91	66.67	76.92
Mamba-2.8B + BERT	C	FISH-ZOOPLANKTON	100.0	53.33	69.57
Mistral-7B + Ada	C	FISH-ZOOPLANKTON	100.0	73.33	84.62
Mistral-7B + BERT	C	FISH-ZOOPLANKTON	100.0	53.33	69.57
Vicuna-7B + Ada	C	FISH-ZOOPLANKTON	100.0	80.0	88.89
Vicuna-7B + BERT	C	FISH-ZOOPLANKTON	100.0	53.33	69.57
GPT-3.5 + Ada	C	ALGAE-ZOOBENTHOS	100.0	33.33	50.0
Falcon-7B + Ada	C	ALGAE-ZOOBENTHOS	77.78	38.89	51.85
Falcon-7B + BERT	C	ALGAE-ZOOBENTHOS	100.0	33.33	50.0
LLaMA-2-7B + Ada	C	ALGAE-ZOOBENTHOS	77.78	38.89	51.85
LLaMA-2-7B + BERT	C	ALGAE-ZOOBENTHOS	100.0	33.33	50.0
MPT-7B + Ada	C	ALGAE-ZOOBENTHOS	77.78	38.89	51.85
MPT-7B + BERT	C	ALGAE-ZOOBENTHOS	100.0	33.33	50.0
Mamba-2.8B + Ada	C	ALGAE-ZOOBENTHOS	77.78	38.89	51.85
Mamba-2.8B + BERT	C	ALGAE-ZOOBENTHOS	83.33	27.78	41.67
Mistral-7B + Ada	C	ALGAE-ZOOBENTHOS	100.0	38.89	56.0
Mistral-7B + BERT	C	ALGAE-ZOOBENTHOS	100.0	33.33	50.0
Vicuna-7B + Ada	C	ALGAE-ZOOBENTHOS	77.78	38.89	51.85
Vicuna-7B + BERT	C	ALGAE-ZOOBENTHOS	100.0	33.33	50.0
GPT-3.5 + Ada	C	TAXR-NCBI (Bacteria)	59.52	100.0	74.63
Falcon-7B + Ada	C	TAXR-NCBI (Bacteria)	58.92	100.0	74.15
Falcon-7B + BERT	C	TAXR-NCBI (Bacteria)	60.14	100.0	75.11
LLaMA-2-7B + Ada	C	TAXR-NCBI (Bacteria)	58.92	100.0	74.15
LLaMA-2-7B + BERT	C	TAXR-NCBI (Bacteria)	60.14	100.0	75.11
MPT-7B + Ada	C	TAXR-NCBI (Bacteria)	58.92	100.0	74.15
MPT-7B + BERT	C	TAXR-NCBI (Bacteria)	60.14	100.0	75.11
Mamba-2.8B + Ada	C	TAXR-NCBI (Bacteria)	58.11	88.0	70.0
Mamba-2.8B + BERT	C	TAXR-NCBI (Bacteria)	60.96	87.43	71.83
Mistral-7B + Ada	C	TAXR-NCBI (Bacteria)	59.32	100.0	74.47
Mistral-7B + BERT	C	TAXR-NCBI (Bacteria)	60.34	100.0	75.27
Vicuna-7B + Ada	C	TAXR-NCBI (Bacteria)	58.97	92.0	71.88
Vicuna-7B + BERT	C	TAXR-NCBI (Bacteria)	62.99	91.43	74.59

Table 11. LLM models results — BIODIV track – Rep is the representation type. Retriever model Top-k is set to 5. PART 2

Model	Rep	Task	Results		
			Prec	Rec	F1
GPT-3.5 + Ada	C	TAXR-NCBI (Chromista)	63.53	98.43	77.22
Falcon-7B + Ada	C	TAXR-NCBI (Chromista)	62.95	98.43	76.79
Falcon-7B + BERT	C	TAXR-NCBI (Chromista)	63.67	98.43	77.33
LLaMA-2-7B + Ada	C	TAXR-NCBI (Chromista)	62.95	98.43	76.79
LLaMA-2-7B + BERT	C	TAXR-NCBI (Chromista)	63.67	98.43	77.33
MPT-7B + Ada	C	TAXR-NCBI (Chromista)	62.95	98.43	76.79
MPT-7B + BERT	C	TAXR-NCBI (Chromista)	63.67	98.43	77.33
Mamba-2.8B + Ada	C	TAXR-NCBI (Chromista)	61.47	83.7	70.89
Mamba-2.8B + BERT	C	TAXR-NCBI (Chromista)	65.04	84.2	73.39
Mistral-7B + Ada	C	TAXR-NCBI (Chromista)	63.35	98.43	77.09
Mistral-7B + BERT	C	TAXR-NCBI (Chromista)	63.82	98.43	77.44
Vicuna-7B + Ada	C	TAXR-NCBI (Chromista)	61.77	85.55	71.74
Vicuna-7B + BERT	C	TAXR-NCBI (Chromista)	63.91	83.56	72.42
GPT-3.5 + Ada	C	TAXR-NCBI (Fungi)	80.58	99.86	89.19
Falcon-7B + Ada	C	TAXR-NCBI (Fungi)	80.46	99.86	89.12
Falcon-7B + BERT	C	TAXR-NCBI (Fungi)	80.75	99.86	89.3
LLaMA-2-7B + Ada	C	TAXR-NCBI (Fungi)	80.46	99.87	89.12
LLaMA-2-7B + BERT	C	TAXR-NCBI (Fungi)	80.75	99.87	89.3
MPT-7B + Ada	C	TAXR-NCBI (Fungi)	80.46	99.87	89.12
MPT-7B + BERT	C	TAXR-NCBI (Fungi)	80.75	99.87	89.3
Mamba-2.8B + Ada	C	TAXR-NCBI (Fungi)	78.89	84.86	81.76
Mamba-2.8B + BERT	C	TAXR-NCBI (Fungi)	80.94	84.52	82.69
Mistral-7B + Ada	C	TAXR-NCBI (Fungi)	80.48	99.87	89.14
Mistral-7B + BERT	C	TAXR-NCBI (Fungi)	80.77	99.86	89.31
Vicuna-7B + Ada	C	TAXR-NCBI (Fungi)	78.31	88.51	83.1
Vicuna-7B + BERT	C	TAXR-NCBI (Fungi)	81.21	88.61	84.75
GPT-3.5 + Ada	C	TAXR-NCBI (Plantae)	76.74	99.21	86.54
Falcon-7B + Ada	C	TAXR-NCBI (Plantae)	76.22	99.22	86.21
Falcon-7B + BERT	C	TAXR-NCBI (Plantae)	76.94	99.09	86.62
LLaMA-2-7B + Ada	C	TAXR-NCBI (Plantae)	76.22	99.24	86.22
LLaMA-2-7B + BERT	C	TAXR-NCBI (Plantae)	76.94	99.1	86.62
MPT-7B + Ada	C	TAXR-NCBI (Plantae)	76.22	99.24	86.22
MPT-7B + BERT	C	TAXR-NCBI (Plantae)	76.94	99.1	86.62
Mamba-2.8B + Ada	C	TAXR-NCBI (Plantae)	74.57	84.33	79.15
Mamba-2.8B + BERT	C	TAXR-NCBI (Plantae)	77.22	84.0	80.47
Mistral-7B + Ada	C	TAXR-NCBI (Plantae)	76.33	99.23	86.29
Mistral-7B + BERT	C	TAXR-NCBI (Plantae)	76.95	99.1	86.63
Vicuna-7B + Ada	C	TAXR-NCBI (Plantae)	74.21	87.82	80.44
Vicuna-7B + BERT	C	TAXR-NCBI (Plantae)	76.85	87.81	81.96
GPT-3.5 + Ada	C	TAXR-NCBI (Protozoa)	75.16	100.0	85.82
Falcon-7B + Ada	C	TAXR-NCBI (Protozoa)	73.76	100.0	84.9
Falcon-7B + BERT	C	TAXR-NCBI (Protozoa)	75.8	100.0	86.23
LLaMA-2-7B + Ada	C	TAXR-NCBI (Protozoa)	73.76	100.0	84.9
LLaMA-2-7B + BERT	C	TAXR-NCBI (Protozoa)	75.8	100.0	86.23
MPT-7B + Ada	C	TAXR-NCBI (Protozoa)	73.76	100.0	84.9
MPT-7B + BERT	C	TAXR-NCBI (Protozoa)	75.8	100.0	86.23
Mamba-2.8B + Ada	C	TAXR-NCBI (Protozoa)	70.6	82.07	75.91
Mamba-2.8B + BERT	C	TAXR-NCBI (Protozoa)	75.57	84.03	79.58
Mistral-7B + Ada	C	TAXR-NCBI (Protozoa)	74.53	100.0	85.41
Mistral-7B + BERT	C	TAXR-NCBI (Protozoa)	75.8	100.0	86.23
Vicuna-7B + Ada	C	TAXR-NCBI (Protozoa)	71.95	87.68	79.04
Vicuna-7B + BERT	C	TAXR-NCBI (Protozoa)	77.27	85.71	81.27

Table 12. LLM models results — BIODIV track – Rep is the representation type. Retriever model Top-k is set to 5. PART 3

Model	Rep	Task	Results		
			Prec	Rec	F1
GPT-3.5 + Ada	CC	ENVO-SWEET	58.29	41.49	48.48
Falcon-7B + Ada	CC	ENVO-SWEET	55.02	53.79	54.4
Falcon-7B + BERT	CC	ENVO-SWEET	81.63	29.81	43.68
LLaMA-2-7B + Ada	CC	ENVO-SWEET	55.02	53.79	54.4
LLaMA-2-7B + BERT	CC	ENVO-SWEET	81.91	29.81	43.72
MPT-7B + Ada	CC	ENVO-SWEET	55.02	53.79	54.4
MPT-7B + BERT	CC	ENVO-SWEET	81.63	29.81	43.68
Mamba-2.8B + Ada	CC	ENVO-SWEET	52.7	48.57	50.55
Mamba-2.8B + BERT	CC	ENVO-SWEET	82.68	26.09	39.66
Mistral-7B + Ada	CC	ENVO-SWEET	65.16	33.91	44.61
Mistral-7B + BERT	CC	ENVO-SWEET	85.41	24.72	38.34
Vicuna-7B + Ada	CC	ENVO-SWEET	47.22	40.12	43.38
Vicuna-7B + BERT	CC	ENVO-SWEET	80.77	20.87	33.17
GPT-3.5 + Ada	CC	FISH-ZOOPLANKTON	100.0	66.67	80.0
Falcon-7B + Ada	CC	FISH-ZOOPLANKTON	100.0	80.0	88.89
Falcon-7B + BERT	CC	FISH-ZOOPLANKTON	100.0	53.33	69.57
LLaMA-2-7B + Ada	CC	FISH-ZOOPLANKTON	100.0	80.0	88.89
LLaMA-2-7B + BERT	CC	FISH-ZOOPLANKTON	100.0	53.33	69.57
MPT-7B + Ada	CC	FISH-ZOOPLANKTON	100.0	80.0	88.89
MPT-7B + BERT	CC	FISH-ZOOPLANKTON	100.0	53.33	69.57
Mamba-2.8B + Ada	CC	FISH-ZOOPLANKTON	100.0	53.33	69.57
Mamba-2.8B + BERT	CC	FISH-ZOOPLANKTON	100.0	53.33	69.57
Mistral-7B + Ada	CC	FISH-ZOOPLANKTON	100.0	33.33	50.0
Mistral-7B + BERT	CC	FISH-ZOOPLANKTON	100.0	33.33	50.0
Vicuna-7B + Ada	CC	FISH-ZOOPLANKTON	91.67	73.33	81.48
Vicuna-7B + BERT	CC	FISH-ZOOPLANKTON	100.0	40.0	57.14
GPT-3.5 + Ada	CC	ALGAE-ZOOBENTHOS	100.0	33.33	50.0
Falcon-7B + Ada	CC	ALGAE-ZOOBENTHOS	77.78	38.89	51.85
Falcon-7B + BERT	CC	ALGAE-ZOOBENTHOS	100.0	33.33	50.0
LLaMA-2-7B + Ada	CC	ALGAE-ZOOBENTHOS	77.78	38.89	51.85
LLaMA-2-7B + BERT	CC	ALGAE-ZOOBENTHOS	100.0	33.33	50.0
MPT-7B + Ada	CC	ALGAE-ZOOBENTHOS	77.78	38.89	51.85
MPT-7B + BERT	CC	ALGAE-ZOOBENTHOS	100.0	33.33	50.0
Mamba-2.8B + Ada	CC	ALGAE-ZOOBENTHOS	77.78	38.89	51.85
Mamba-2.8B + BERT	CC	ALGAE-ZOOBENTHOS	100.0	22.22	36.36
Mistral-7B + Ada	CC	ALGAE-ZOOBENTHOS	100.0	22.22	36.36
Mistral-7B + BERT	CC	ALGAE-ZOOBENTHOS	100.0	16.67	28.57
Vicuna-7B + Ada	CC	ALGAE-ZOOBENTHOS	85.71	33.33	48.0
Vicuna-7B + BERT	CC	ALGAE-ZOOBENTHOS	100.0	33.33	50.0
GPT-3.5 + Ada	CC	TAXR-NCBI (Bacteria)	59.73	100.0	74.79
Falcon-7B + Ada	CC	TAXR-NCBI (Bacteria)	59.12	100.0	74.31
Falcon-7B + BERT	CC	TAXR-NCBI (Bacteria)	60.14	100.0	75.11
LLaMA-2-7B + Ada	CC	TAXR-NCBI (Bacteria)	58.92	100.0	74.15
LLaMA-2-7B + BERT	CC	TAXR-NCBI (Bacteria)	60.14	100.0	75.11
MPT-7B + Ada	CC	TAXR-NCBI (Bacteria)	58.92	100.0	74.15
MPT-7B + BERT	CC	TAXR-NCBI (Bacteria)	60.14	100.0	75.11
Mamba-2.8B + Ada	CC	TAXR-NCBI (Bacteria)	56.34	86.29	68.17
Mamba-2.8B + BERT	CC	TAXR-NCBI (Bacteria)	60.55	88.57	71.93
Mistral-7B + Ada	CC	TAXR-NCBI (Bacteria)	60.34	100.0	75.27
Mistral-7B + BERT	CC	TAXR-NCBI (Bacteria)	61.19	100.0	75.92
Vicuna-7B + Ada	CC	TAXR-NCBI (Bacteria)	59.32	100.0	74.47
Vicuna-7B + BERT	CC	TAXR-NCBI (Bacteria)	60.34	100.0	75.27

Table 13. LLM models results — BIODIV track – Rep is the representation type. Retriever model Top-k is set to 5. PART 4

Model	Rep	Task	Results		
			Prec	Rec	F1
GPT-3.5 + Ada	CC	TAXR-NCBI (Chromista)	63.94	98.43	77.52
Falcon-7B + Ada	CC	TAXR-NCBI (Chromista)	63.09	98.43	76.9
Falcon-7B + BERT	CC	TAXR-NCBI (Chromista)	63.85	98.43	77.46
LLaMA-2-7B + Ada	CC	TAXR-NCBI (Chromista)	63.01	98.43	76.83
LLaMA-2-7B + BERT	CC	TAXR-NCBI (Chromista)	63.73	98.43	77.37
MPT-7B + Ada	CC	TAXR-NCBI (Chromista)	63.01	98.43	76.83
MPT-7B + BERT	CC	TAXR-NCBI (Chromista)	63.73	98.43	77.37
Mamba-2.8B + Ada	CC	TAXR-NCBI (Chromista)	61.35	82.92	70.52
Mamba-2.8B + BERT	CC	TAXR-NCBI (Chromista)	63.81	83.2	72.23
Mistral-7B + Ada	CC	TAXR-NCBI (Chromista)	63.85	98.43	77.46
Mistral-7B + BERT	CC	TAXR-NCBI (Chromista)	64.58	98.36	77.97
Vicuna-7B + Ada	CC	TAXR-NCBI (Chromista)	63.69	98.36	77.31
Vicuna-7B + BERT	CC	TAXR-NCBI (Chromista)	64.63	98.43	78.03
GPT-3.5 + Ada	CC	TAXR-NCBI (Fungi)	80.71	99.85	89.27
Falcon-7B + Ada	CC	TAXR-NCBI (Fungi)	80.82	99.87	89.34
Falcon-7B + BERT	CC	TAXR-NCBI (Fungi)	81.09	99.87	89.51
LLaMA-2-7B + Ada	CC	TAXR-NCBI (Fungi)	80.75	99.87	89.3
LLaMA-2-7B + BERT	CC	TAXR-NCBI (Fungi)	81.06	99.87	89.49
MPT-7B + Ada	CC	TAXR-NCBI (Fungi)	80.59	99.87	89.2
MPT-7B + BERT	CC	TAXR-NCBI (Fungi)	80.89	99.87	89.39
Mamba-2.8B + Ada	CC	TAXR-NCBI (Fungi)	78.89	85.95	82.27
Mamba-2.8B + BERT	CC	TAXR-NCBI (Fungi)	80.86	85.83	83.27
Mistral-7B + Ada	CC	TAXR-NCBI (Fungi)	81.29	99.77	89.59
Mistral-7B + BERT	CC	TAXR-NCBI (Fungi)	81.52	99.78	89.73
Vicuna-7B + Ada	CC	TAXR-NCBI (Fungi)	81.63	99.79	89.8
Vicuna-7B + BERT	CC	TAXR-NCBI (Fungi)	82.02	99.71	90.01
GPT-3.5 + Ada	CC	TAXR-NCBI (Plantae)	76.55	99.19	86.41
Falcon-7B + Ada	CC	TAXR-NCBI (Plantae)	76.53	99.16	86.39
Falcon-7B + BERT	CC	TAXR-NCBI (Plantae)	77.25	99.03	86.79
LLaMA-2-7B + Ada	CC	TAXR-NCBI (Plantae)	76.4	99.24	86.34
LLaMA-2-7B + BERT	CC	TAXR-NCBI (Plantae)	77.13	99.1	86.74
MPT-7B + Ada	CC	TAXR-NCBI (Plantae)	76.34	99.24	86.3
MPT-7B + BERT	CC	TAXR-NCBI (Plantae)	77.05	99.1	86.7
Mamba-2.8B + Ada	CC	TAXR-NCBI (Plantae)	74.48	84.35	79.11
Mamba-2.8B + BERT	CC	TAXR-NCBI (Plantae)	76.79	84.29	80.36
Mistral-7B + Ada	CC	TAXR-NCBI (Plantae)	77.28	98.97	86.79
Mistral-7B + BERT	CC	TAXR-NCBI (Plantae)	77.98	98.92	87.21
Vicuna-7B + Ada	CC	TAXR-NCBI (Plantae)	77.21	99.04	86.77
Vicuna-7B + BERT	CC	TAXR-NCBI (Plantae)	78.03	98.82	87.2
GPT-3.5 + Ada	CC	TAXR-NCBI (Protozoa)	75.96	100.0	86.34
Falcon-7B + Ada	CC	TAXR-NCBI (Protozoa)	73.91	100.0	85.0
Falcon-7B + BERT	CC	TAXR-NCBI (Protozoa)	75.96	100.0	86.34
LLaMA-2-7B + Ada	CC	TAXR-NCBI (Protozoa)	74.07	100.0	85.1
LLaMA-2-7B + BERT	CC	TAXR-NCBI (Protozoa)	76.12	100.0	86.44
MPT-7B + Ada	CC	TAXR-NCBI (Protozoa)	73.91	100.0	85.0
MPT-7B + BERT	CC	TAXR-NCBI (Protozoa)	75.96	100.0	86.34
Mamba-2.8B + Ada	CC	TAXR-NCBI (Protozoa)	72.17	85.71	78.36
Mamba-2.8B + BERT	CC	TAXR-NCBI (Protozoa)	76.08	83.75	79.73
Mistral-7B + Ada	CC	TAXR-NCBI (Protozoa)	75.96	100.0	86.34
Mistral-7B + BERT	CC	TAXR-NCBI (Protozoa)	77.11	100.0	87.07
Vicuna-7B + Ada	CC	TAXR-NCBI (Protozoa)	74.79	99.72	85.47
Vicuna-7B + BERT	CC	TAXR-NCBI (Protozoa)	77.17	99.44	86.9

Table 14. LLM models results — BIODIV track – Rep is the representation type. Retriever model Top-k is set to 5. PART 5

Model	Rep	Task	Results		
			Prec	Rec	F1
GPT-3.5 + Ada	CP	ENVO-SWEET	66.59	36.65	47.28
Falcon-7B + Ada	CP	ENVO-SWEET	55.02	53.79	54.4
Falcon-7B + BERT	CP	ENVO-SWEET	81.63	29.81	43.68
LLaMA-2-7B + Ada	CP	ENVO-SWEET	55.02	53.79	54.4
LLaMA-2-7B + BERT	CP	ENVO-SWEET	81.63	29.81	43.68
MPT-7B + Ada	CP	ENVO-SWEET	55.02	53.79	54.4
MPT-7B + BERT	CP	ENVO-SWEET	81.63	29.81	43.68
Mamba-2.8B + Ada	CP	ENVO-SWEET	49.73	45.34	47.43
Mamba-2.8B + BERT	CP	ENVO-SWEET	80.57	24.72	37.83
Mistral-7B + Ada	CP	ENVO-SWEET	75.38	31.18	44.11
Mistral-7B + BERT	CP	ENVO-SWEET	87.45	25.09	39.0
Vicuna-7B + Ada	CP	ENVO-SWEET	46.97	37.52	41.71
Vicuna-7B + BERT	CP	ENVO-SWEET	82.59	20.62	33.0
GPT-3.5 + Ada	CP	FISH-ZOOPLANKTON	100.0	66.67	80.0
Falcon-7B + Ada	CP	FISH-ZOOPLANKTON	100.0	80.0	88.89
Falcon-7B + BERT	CP	FISH-ZOOPLANKTON	100.0	53.33	69.57
LLaMA-2-7B + Ada	CP	FISH-ZOOPLANKTON	100.0	80.0	88.89
LLaMA-2-7B + BERT	CP	FISH-ZOOPLANKTON	100.0	53.33	69.57
MPT-7B + Ada	CP	FISH-ZOOPLANKTON	100.0	80.0	88.89
MPT-7B + BERT	CP	FISH-ZOOPLANKTON	100.0	53.33	69.57
Mamba-2.8B + Ada	CP	FISH-ZOOPLANKTON	87.5	46.67	60.87
Mamba-2.8B + BERT	CP	FISH-ZOOPLANKTON	100.0	40.0	57.14
Mistral-7B + Ada	CP	FISH-ZOOPLANKTON	100.0	46.67	63.64
Mistral-7B + BERT	CP	FISH-ZOOPLANKTON	100.0	40.0	57.14
Vicuna-7B + Ada	CP	FISH-ZOOPLANKTON	90.91	66.67	76.92
Vicuna-7B + BERT	CP	FISH-ZOOPLANKTON	100.0	40.0	57.14
GPT-3.5 + Ada	CP	ALGAE-ZOOBENTHOS	100.0	33.33	50.0
Falcon-7B + Ada	CP	ALGAE-ZOOBENTHOS	77.78	38.89	51.85
Falcon-7B + BERT	CP	ALGAE-ZOOBENTHOS	100.0	33.33	50.0
LLaMA-2-7B + Ada	CP	ALGAE-ZOOBENTHOS	77.78	38.89	51.85
LLaMA-2-7B + BERT	CP	ALGAE-ZOOBENTHOS	100.0	33.33	50.0
MPT-7B + Ada	CP	ALGAE-ZOOBENTHOS	77.78	38.89	51.85
MPT-7B + BERT	CP	ALGAE-ZOOBENTHOS	100.0	33.33	50.0
Mamba-2.8B + Ada	CP	ALGAE-ZOOBENTHOS	77.78	38.89	51.85
Mamba-2.8B + BERT	CP	ALGAE-ZOOBENTHOS	100.0	27.78	43.48
Mistral-7B + Ada	CP	ALGAE-ZOOBENTHOS	100.0	27.78	43.48
Mistral-7B + BERT	CP	ALGAE-ZOOBENTHOS	100.0	22.22	36.36
Vicuna-7B + Ada	CP	ALGAE-ZOOBENTHOS	87.5	38.89	53.85
Vicuna-7B + BERT	CP	ALGAE-ZOOBENTHOS	100.0	33.33	50.0
GPT-3.5 + Ada	CP	TAXR-NCBI (Bacteria)	67.97	99.43	80.74
Falcon-7B + Ada	CP	TAXR-NCBI (Bacteria)	58.92	100.0	74.15
Falcon-7B + BERT	CP	TAXR-NCBI (Bacteria)	60.14	100.0	75.11
LLaMA-2-7B + Ada	CP	TAXR-NCBI (Bacteria)	58.92	100.0	74.15
LLaMA-2-7B + BERT	CP	TAXR-NCBI (Bacteria)	60.14	100.0	75.11
MPT-7B + Ada	CP	TAXR-NCBI (Bacteria)	58.92	100.0	74.15
MPT-7B + BERT	CP	TAXR-NCBI (Bacteria)	60.14	100.0	75.11
Mamba-2.8B + Ada	CP	TAXR-NCBI (Bacteria)	58.3	86.29	69.59
Mamba-2.8B + BERT	CP	TAXR-NCBI (Bacteria)	58.87	83.43	69.03
Mistral-7B + Ada	CP	TAXR-NCBI (Bacteria)	60.84	99.43	75.49
Mistral-7B + BERT	CP	TAXR-NCBI (Bacteria)	61.27	99.43	75.82
Vicuna-7B + Ada	CP	TAXR-NCBI (Bacteria)	59.18	99.43	74.2
Vicuna-7B + BERT	CP	TAXR-NCBI (Bacteria)	60.55	100.0	75.43

Table 15. LLM models results — BIODIV track – Rep is the representation type. Retriever model Top-k is set to 5. PART 6

Model	Rep	Task	Results		
			Prec	Rec	F1
GPT-3.5 + Ada	CP	TAXR-NCBI (Chromista)	69.88	98.08	81.61
Falcon-7B + Ada	CP	TAXR-NCBI (Chromista)	62.95	98.43	76.79
Falcon-7B + BERT	CP	TAXR-NCBI (Chromista)	63.67	98.43	77.33
LLaMA-2-7B + Ada	CP	TAXR-NCBI (Chromista)	62.95	98.43	76.79
LLaMA-2-7B + BERT	CP	TAXR-NCBI (Chromista)	63.67	98.43	77.33
MPT-7B + Ada	CP	TAXR-NCBI (Chromista)	62.95	98.43	76.79
MPT-7B + BERT	CP	TAXR-NCBI (Chromista)	63.67	98.43	77.33
Mamba-2.8B + Ada	CP	TAXR-NCBI (Chromista)	61.66	84.13	71.16
Mamba-2.8B + BERT	CP	TAXR-NCBI (Chromista)	63.68	84.84	72.75
Mistral-7B + Ada	CP	TAXR-NCBI (Chromista)	64.29	98.29	77.74
Mistral-7B + BERT	CP	TAXR-NCBI (Chromista)	64.55	98.22	77.9
Vicuna-7B + Ada	CP	TAXR-NCBI (Chromista)	63.22	98.36	76.97
Vicuna-7B + BERT	CP	TAXR-NCBI (Chromista)	64.0	98.43	77.57
GPT-3.5 + Ada	CP	TAXR-NCBI (Fungi)	86.97	99.08	92.64
Falcon-7B + Ada	CP	TAXR-NCBI (Fungi)	80.46	99.87	89.12
Falcon-7B + BERT	CP	TAXR-NCBI (Fungi)	80.76	99.86	89.3
LLaMA-2-7B + Ada	CP	TAXR-NCBI (Fungi)	80.46	99.87	89.12
LLaMA-2-7B + BERT	CP	TAXR-NCBI (Fungi)	80.75	99.87	89.3
MPT-7B + Ada	CP	TAXR-NCBI (Fungi)	80.46	99.87	89.12
MPT-7B + BERT	CP	TAXR-NCBI (Fungi)	80.75	99.87	89.3
Mamba-2.8B + Ada	CP	TAXR-NCBI (Fungi)	78.7	86.4	82.37
Mamba-2.8B + BERT	CP	TAXR-NCBI (Fungi)	80.58	85.95	83.18
Mistral-7B + Ada	CP	TAXR-NCBI (Fungi)	81.79	99.72	89.87
Mistral-7B + BERT	CP	TAXR-NCBI (Fungi)	82.05	99.74	90.03
Vicuna-7B + Ada	CP	TAXR-NCBI (Fungi)	80.89	99.75	89.34
Vicuna-7B + BERT	CP	TAXR-NCBI (Fungi)	81.3	99.74	89.58
GPT-3.5 + Ada	CP	TAXR-NCBI (Plantae)	82.6	96.35	88.95
Falcon-7B + Ada	CP	TAXR-NCBI (Plantae)	76.24	99.19	86.21
Falcon-7B + BERT	CP	TAXR-NCBI (Plantae)	76.95	99.07	86.62
LLaMA-2-7B + Ada	CP	TAXR-NCBI (Plantae)	76.22	99.24	86.22
LLaMA-2-7B + BERT	CP	TAXR-NCBI (Plantae)	76.94	99.1	86.62
MPT-7B + Ada	CP	TAXR-NCBI (Plantae)	76.22	99.24	86.22
MPT-7B + BERT	CP	TAXR-NCBI (Plantae)	76.94	99.1	86.62
Mamba-2.8B + Ada	CP	TAXR-NCBI (Plantae)	74.71	84.9	79.48
Mamba-2.8B + BERT	CP	TAXR-NCBI (Plantae)	77.41	85.33	81.18
Mistral-7B + Ada	CP	TAXR-NCBI (Plantae)	76.98	99.14	86.66
Mistral-7B + BERT	CP	TAXR-NCBI (Plantae)	77.41	99.02	86.89
Vicuna-7B + Ada	CP	TAXR-NCBI (Plantae)	76.85	98.36	86.29
Vicuna-7B + BERT	CP	TAXR-NCBI (Plantae)	77.76	98.26	86.82
GPT-3.5 + Ada	CP	TAXR-NCBI (Protozoa)	86.06	98.6	91.91
Falcon-7B + Ada	CP	TAXR-NCBI (Protozoa)	73.76	100.0	84.9
Falcon-7B + BERT	CP	TAXR-NCBI (Protozoa)	75.8	100.0	86.23
LLaMA-2-7B + Ada	CP	TAXR-NCBI (Protozoa)	73.76	100.0	84.9
LLaMA-2-7B + BERT	CP	TAXR-NCBI (Protozoa)	75.8	100.0	86.23
MPT-7B + Ada	CP	TAXR-NCBI (Protozoa)	73.76	100.0	84.9
MPT-7B + BERT	CP	TAXR-NCBI (Protozoa)	75.8	100.0	86.23
Mamba-2.8B + Ada	CP	TAXR-NCBI (Protozoa)	70.69	83.75	76.67
Mamba-2.8B + BERT	CP	TAXR-NCBI (Protozoa)	73.89	84.03	78.64
Mistral-7B + Ada	CP	TAXR-NCBI (Protozoa)	75.96	100.0	86.34
Mistral-7B + BERT	CP	TAXR-NCBI (Protozoa)	76.28	100.0	86.55
Vicuna-7B + Ada	CP	TAXR-NCBI (Protozoa)	74.01	99.72	84.96
Vicuna-7B + BERT	CP	TAXR-NCBI (Protozoa)	76.07	99.72	86.3

Table 16. LLM models results — PHENOTYPE track – Rep is the representation type. Retriever model Top-k is set to 5. PART 1

Model	**Rep**	**Task**	**Results**		
			Prec	Rec	F1
GPT-3.5 + Ada	C	DOID-ORDO	45.98	99.35	62.86
Falcon-7B + Ada	C	DOID-ORDO	39.99	99.27	57.01
Falcon-7B + BERT	C	DOID-ORDO	73.01	98.63	83.91
LLaMA-2-7B + Ada	C	DOID-ORDO	40.01	99.35	57.04
LLaMA-2-7B + BERT	C	DOID-ORDO	73.01	98.63	83.91
MPT-7B + Ada	C	DOID-ORDO	40.01	99.35	57.04
MPT-7B + BERT	C	DOID-ORDO	73.01	98.63	83.91
Mamba-2.8B + Ada	C	DOID-ORDO	35.83	85.13	50.43
Mamba-2.8B + BERT	C	DOID-ORDO	70.19	83.19	76.14
Mistral-7B + Ada	C	DOID-ORDO	42.92	99.27	59.93
Mistral-7B + BERT	C	DOID-ORDO	74.42	98.54	84.8
Vicuna-7B + Ada	C	DOID-ORDO	37.14	90.14	52.61
Vicuna-7B + BERT	C	DOID-ORDO	71.18	89.25	79.2
GPT-3.5 + Ada	C	HP-MP	28.92	99.14	44.78
Falcon-7B + Ada	C	HP-MP	24.65	99.14	39.48
Falcon-7B + BERT	C	HP-MP	54.7	98.56	70.36
LLaMA-2-7B + Ada	C	HP-MP	24.69	99.28	39.54
LLaMA-2-7B + BERT	C	HP-MP	54.78	98.85	70.49
MPT-7B + Ada	C	HP-MP	24.69	99.28	39.54
MPT-7B + BERT	C	HP-MP	54.78	98.85	70.49
Mamba-2.8B + Ada	C	HP-MP	19.51	77.44	31.17
Mamba-2.8B + BERT	C	HP-MP	53.1	82.33	64.56
Mistral-7B + Ada	C	HP-MP	27.0	99.14	42.44
Mistral-7B + BERT	C	HP-MP	56.17	98.71	71.6
Vicuna-7B + Ada	C	HP-MP	21.22	83.76	33.86
Vicuna-7B + BERT	C	HP-MP	53.98	82.76	65.34

Table 17. LLM models results — PHENOTYPE track – Rep is the representation type. Retriever model Top-k is set to 5. PART 2

Model	**Rep**	**Task**	**Results**		
			Prec	Rec	F1
GPT-3.5 + Ada	CC	DOID-ORDO	46.37	99.19	63.2
Falcon-7B + Ada	CC	DOID-ORDO	39.94	99.19	56.95
Falcon-7B + BERT	CC	DOID-ORDO	72.93	98.46	83.8
LLaMA-2-7B + Ada	CC	DOID-ORDO	39.78	98.71	56.71
LLaMA-2-7B + BERT	CC	DOID-ORDO	72.73	97.9	83.46
MPT-7B + Ada	CC	DOID-ORDO	39.99	99.27	57.01
MPT-7B + BERT	CC	DOID-ORDO	72.99	98.54	83.87
Mamba-2.8B + Ada	CC	DOID-ORDO	34.86	83.1	49.12
Mamba-2.8B + BERT	CC	DOID-ORDO	70.41	82.7	76.06
Mistral-7B + Ada	CC	DOID-ORDO	49.29	89.98	63.69
Mistral-7B + BERT	CC	DOID-ORDO	78.34	89.17	83.4
Vicuna-7B + Ada	CC	DOID-ORDO	37.69	91.43	53.37
Vicuna-7B + BERT	CC	DOID-ORDO	71.33	90.7	79.86
GPT-3.5 + Ada	CC	HP-MP	28.95	98.13	44.71
Falcon-7B + Ada	CC	HP-MP	24.6	98.99	39.41
Falcon-7B + BERT	CC	HP-MP	54.75	98.56	70.4
LLaMA-2-7B + Ada	CC	HP-MP	24.65	99.14	39.48
LLaMA-2-7B + BERT	CC	HP-MP	54.86	98.85	70.56
MPT-7B + Ada	CC	HP-MP	24.69	99.28	39.54
MPT-7B + BERT	CC	HP-MP	54.78	98.85	70.49
Mamba-2.8B + Ada	CC	HP-MP	20.9	81.9	33.3
Mamba-2.8B + BERT	CC	HP-MP	53.83	84.91	65.89
Mistral-7B + Ada	CC	HP-MP	43.84	73.56	54.94
Mistral-7B + BERT	CC	HP-MP	82.16	72.13	76.82
Vicuna-7B + Ada	CC	HP-MP	21.85	83.91	34.67
Vicuna-7B + BERT	CC	HP-MP	55.28	84.2	66.74

Table 18. LLM models results — PHENOTYPE track – Rep is the representation type. Retriever model Top-k is set to 5. PART 3

Model	**Rep**	**Task**	**Results**		
			Prec	Rec	F1
GPT-3.5 + Ada	CP	DOID-ORDO	55.39	94.75	69.91
Falcon-7B + Ada	CP	DOID-ORDO	39.94	99.19	56.95
Falcon-7B + BERT	CP	DOID-ORDO	72.93	98.46	83.8
LLaMA-2-7B + Ada	CP	DOID-ORDO	40.01	99.35	57.04
LLaMA-2-7B + BERT	CP	DOID-ORDO	73.01	98.63	83.91
MPT-7B + Ada	CP	DOID-ORDO	40.01	99.35	57.04
MPT-7B + BERT	CP	DOID-ORDO	73.01	98.63	83.91
Mamba-2.8B + Ada	CP	DOID-ORDO	36.66	87.07	51.59
Mamba-2.8B + BERT	CP	DOID-ORDO	71.43	86.5	78.24
Mistral-7B + Ada	CP	DOID-ORDO	63.53	94.34	75.93
Mistral-7B + BERT	CP	DOID-ORDO	85.8	94.26	89.83
Vicuna-7B + Ada	CP	DOID-ORDO	39.58	95.07	55.89
Vicuna-7B + BERT	CP	DOID-ORDO	72.75	95.39	82.55
GPT-3.5 + Ada	CP	HP-MP	29.08	96.7	44.72
Falcon-7B + Ada	CP	HP-MP	24.57	98.56	39.33
Falcon-7B + BERT	CP	HP-MP	54.57	97.84	70.06
LLaMA-2-7B + Ada	CP	HP-MP	24.69	99.28	39.54
LLaMA-2-7B + BERT	CP	HP-MP	54.78	98.85	70.49
MPT-7B + Ada	CP	HP-MP	24.69	99.28	39.54
MPT-7B + BERT	CP	HP-MP	54.78	98.85	70.49
Mamba-2.8B + Ada	CP	HP-MP	20.32	79.89	32.4
Mamba-2.8B + BERT	CP	HP-MP	51.85	76.58	61.83
Mistral-7B + Ada	CP	HP-MP	53.73	95.26	68.7
Mistral-7B + BERT	CP	HP-MP	76.67	95.4	85.02
Vicuna-7B + Ada	CP	HP-MP	23.62	85.63	37.03
Vicuna-7B + BERT	CP	HP-MP	57.09	84.48	68.13

Table 19. LLM models results — CommonKG track – Rep is the representation type. Retriever model Top-k is set to 5. PART 1

Model	**Rep**	**Task**	**Results**		
			Prec	Rec	F1
GPT-3.5 + Ada	C	Nell-DBpedia	100.0	89.15	94.26
Falcon-7B + Ada	C	Nell-DBpedia	98.29	89.15	93.5
Falcon-7B + BERT	C	Nell-DBpedia	100.0	79.07	88.31
LLaMA-2-7B + Ada	C	Nell-DBpedia	98.29	89.15	93.5
LLaMA-2-7B + BERT	C	Nell-DBpedia	100.0	79.07	88.31
MPT-7B + Ada	C	Nell-DBpedia	98.29	89.15	93.5
MPT-7B + BERT	C	Nell-DBpedia	100.0	79.07	88.31
Mamba-2.8B + Ada	C	Nell-DBpedia	97.09	77.52	86.21
Mamba-2.8B + BERT	C	Nell-DBpedia	100.0	65.89	79.44
Mistral-7B + Ada	C	Nell-DBpedia	100.0	87.6	93.39
Mistral-7B + BERT	C	Nell-DBpedia	100.0	79.07	88.31
Vicuna-7B + Ada	C	Nell-DBpedia	98.26	87.6	92.62
Vicuna-7B + BERT	C	Nell-DBpedia	100.0	78.29	87.83
GPT-3.5 + Ada	C	Yago-Wikidata	100.0	82.57	90.45
Falcon-7B + Ada	C	Yago-Wikidata	100.0	85.53	92.2
Falcon-7B + BERT	C	Yago-Wikidata	100.0	46.71	63.68
LLaMA-2-7B + Ada	C	Yago-Wikidata	100.0	85.53	92.2
LLaMA-2-7B + BERT	C	Yago-Wikidata	100.0	47.04	63.98
MPT-7B + Ada	C	Yago-Wikidata	100.0	85.53	92.2
MPT-7B + BERT	C	Yago-Wikidata	100.0	47.04	63.98
Mamba-2.8B + Ada	C	Yago-Wikidata	98.61	70.07	81.92
Mamba-2.8B + BERT	C	Yago-Wikidata	100.0	40.46	57.61
Mistral-7B + Ada	C	Yago-Wikidata	100.0	81.58	89.86
Mistral-7B + BERT	C	Yago-Wikidata	100.0	46.71	63.68
Vicuna-7B + Ada	C	Yago-Wikidata	99.61	84.54	91.46
Vicuna-7B + BERT	C	Yago-Wikidata	100.0	46.05	63.06

Table 20. LLM models results — CommonKG track – Rep is the representation type. Retriever model Top-k is set to 5. PART 2

Model	Rep	Task	Results		
			Prec	Rec	F1
GPT-3.5 + Ada	CC	Nell-DBpedia	100.0	88.37	93.83
Falcon-7B + Ada	CC	Nell-DBpedia	98.29	89.15	93.5
Falcon-7B + BERT	CC	Nell-DBpedia	100.0	79.07	88.31
LLaMA-2-7B + Ada	CC	Nell-DBpedia	98.29	89.15	93.5
LLaMA-2-7B + BERT	CC	Nell-DBpedia	100.0	79.07	88.31
MPT-7B + Ada	CC	Nell-DBpedia	98.29	89.15	93.5
MPT-7B + BERT	CC	Nell-DBpedia	100.0	79.07	88.31
Mamba-2.8B + Ada	CC	Nell-DBpedia	94.17	75.19	83.62
Mamba-2.8B + BERT	CC	Nell-DBpedia	100.0	66.67	80.0
Mistral-7B + Ada	CC	Nell-DBpedia	99.11	86.05	92.12
Mistral-7B + BERT	CC	Nell-DBpedia	100.0	78.29	87.83
Vicuna-7B + Ada	CC	Nell-DBpedia	98.18	83.72	90.38
Vicuna-7B + BERT	CC	Nell-DBpedia	100.0	72.87	84.3
GPT-3.5 + Ada	CC	Yago-Wikidata	100.0	85.2	92.01
Falcon-7B + Ada	CC	Yago-Wikidata	100.0	85.2	92.01
Falcon-7B + BERT	CC	Yago-Wikidata	100.0	47.04	63.98
LLaMA-2-7B + Ada	CC	Yago-Wikidata	100.0	85.53	92.2
LLaMA-2-7B + BERT	CC	Yago-Wikidata	100.0	47.04	63.98
MPT-7B + Ada	CC	Yago-Wikidata	100.0	85.2	92.01
MPT-7B + BERT	CC	Yago-Wikidata	100.0	47.04	63.98
Mamba-2.8B + Ada	CC	Yago-Wikidata	99.09	71.38	82.98
Mamba-2.8B + BERT	CC	Yago-Wikidata	100.0	38.49	55.58
Mistral-7B + Ada	CC	Yago-Wikidata	100.0	82.24	90.25
Mistral-7B + BERT	CC	Yago-Wikidata	100.0	45.72	62.75
Vicuna-7B + Ada	CC	Yago-Wikidata	100.0	83.55	91.04
Vicuna-7B + BERT	CC	Yago-Wikidata	100.0	44.41	61.5

Table 21. LLM models results — CommonKG track – Rep is the representation type. Retriever model Top-k is set to 5. PART 3

Model	Rep	Task	Results		
			Prec	Rec	F1
GPT-3.5 + Ada	CP	Nell-DBpedia	100.0	88.37	93.83
Falcon-7B + Ada	CP	Nell-DBpedia	98.29	89.15	93.5
Falcon-7B + BERT	CP	Nell-DBpedia	100.0	79.07	88.31
LLaMA-2-7B + Ada	CP	Nell-DBpedia	98.29	89.15	93.5
LLaMA-2-7B + BERT	CP	Nell-DBpedia	100.0	79.07	88.31
MPT-7B + Ada	CP	Nell-DBpedia	98.29	89.15	93.5
MPT-7B + BERT	CP	Nell-DBpedia	100.0	79.07	88.31
Mamba-2.8B + Ada	CP	Nell-DBpedia	96.91	72.87	83.19
Mamba-2.8B + BERT	CP	Nell-DBpedia	100.0	69.77	82.19
Mistral-7B + Ada	CP	Nell-DBpedia	98.26	87.6	92.62
Mistral-7B + BERT	CP	Nell-DBpedia	100.0	79.07	88.31
Vicuna-7B + Ada	CP	Nell-DBpedia	98.29	89.15	93.5
Vicuna-7B + BERT	CP	Nell-DBpedia	100.0	79.07	88.31
GPT-3.5 + Ada	CP	Yago-Wikidata	100.0	85.53	92.2
Falcon-7B + Ada	CP	Yago-Wikidata	100.0	85.53	92.2
Falcon-7B + BERT	CP	Yago-Wikidata	100.0	47.04	63.98
LLaMA-2-7B + Ada	CP	Yago-Wikidata	100.0	85.53	92.2
LLaMA-2-7B + BERT	CP	Yago-Wikidata	100.0	47.04	63.98
MPT-7B + Ada	CP	Yago-Wikidata	100.0	85.53	92.2
MPT-7B + BERT	CP	Yago-Wikidata	100.0	47.04	63.98
Mamba-2.8B + Ada	CP	Yago-Wikidata	100.0	75.0	85.71
Mamba-2.8B + BERT	CP	Yago-Wikidata	100.0	40.13	57.28
Mistral-7B + Ada	CP	Yago-Wikidata	100.0	83.88	91.23
Mistral-7B + BERT	CP	Yago-Wikidata	100.0	46.71	63.68
Vicuna-7B + Ada	CP	Yago-Wikidata	100.0	85.2	92.01
Vicuna-7B + BERT	CP	Yago-Wikidata	100.0	47.37	64.29

Table 22. LLM models results — BIO-ML track – Rep is the representation type. Retriever model Top-k is set to 5. PART 1

Model	Rep	Task	Results		
			Prec	Rec	F1
GPT-3.5 + Ada	C	NCIT-DOID(disease)	86.19	80.07	83.02
Falcon-7B + Ada	C	NCIT-DOID(disease)	83.05	81.03	82.03
Falcon-7B + BERT	C	NCIT-DOID(disease)	93.9	72.24	81.65
LLaMA-2-7B + Ada	C	NCIT-DOID(disease)	83.07	81.07	82.06
LLaMA-2-7B + BERT	C	NCIT-DOID(disease)	93.9	72.28	81.68
MPT-7B + Ada	C	NCIT-DOID(disease)	83.07	81.07	82.06
MPT-7B + BERT	C	NCIT-DOID(disease)	93.9	72.28	81.68
Mamba-2.8B + Ada	C	NCIT-DOID(disease)	74.58	67.99	71.13
Mamba-2.8B + BERT	C	NCIT-DOID(disease)	89.39	59.13	71.18
Mistral-7B + Ada	C	NCIT-DOID(disease)	83.84	79.26	81.48
Mistral-7B + BERT	C	NCIT-DOID(disease)	94.12	70.7	80.75
Vicuna-7B + Ada	C	NCIT-DOID(disease)	79.0	74.43	76.65
Vicuna-7B + BERT	C	NCIT-DOID(disease)	91.95	66.09	76.91
GPT-3.5 + Ada	C	OMIM-ORDO(disease)	72.03	57.73	64.09
Falcon-7B + Ada	C	OMIM-ORDO(disease)	69.15	58.61	63.45
Falcon-7B + BERT	C	OMIM-ORDO(disease)	88.79	43.4	58.3
LLaMA-2-7B + Ada	C	OMIM-ORDO(disease)	69.18	58.64	63.48
LLaMA-2-7B + BERT	C	OMIM-ORDO(disease)	88.74	43.4	58.29
MPT-7B + Ada	C	OMIM-ORDO(disease)	69.18	58.64	63.48
MPT-7B + BERT	C	OMIM-ORDO(disease)	88.74	43.4	58.29
Mamba-2.8B + Ada	C	OMIM-ORDO(disease)	63.33	51.14	56.59
Mamba-2.8B + BERT	C	OMIM-ORDO(disease)	85.31	37.14	51.75
Mistral-7B + Ada	C	OMIM-ORDO(disease)	70.26	58.16	63.64
Mistral-7B + BERT	C	OMIM-ORDO(disease)	88.93	42.97	57.95
Vicuna-7B + Ada	C	OMIM-ORDO(disease)	67.29	56.06	61.16
Vicuna-7B + BERT	C	OMIM-ORDO(disease)	87.49	40.42	55.29
GPT-3.5 + Ada	C	SNOMED-FMA(body)	18.04	30.17	22.58
Falcon-7B + Ada	C	SNOMED-FMA(body)	16.93	28.98	21.38
Falcon-7B + BERT	C	SNOMED-FMA(body)	54.67	5.97	10.76
LLaMA-2-7B + Ada	C	SNOMED-FMA(body)	16.94	29.0	21.38
LLaMA-2-7B + BERT	C	SNOMED-FMA(body)	54.6	5.97	10.76
MPT-7B + Ada	C	SNOMED-FMA(body)	16.94	29.0	21.38
MPT-7B + BERT	C	SNOMED-FMA(body)	54.6	5.97	10.76
Mamba-2.8B + Ada	C	SNOMED-FMA(body)	16.66	28.8	21.11
Mamba-2.8B + BERT	C	SNOMED-FMA(body)	53.19	5.28	9.6
Mistral-7B + Ada	C	SNOMED-FMA(body)	17.93	29.41	22.28
Mistral-7B + BERT	C	SNOMED-FMA(body)	55.48	5.93	10.71
Vicuna-7B + Ada	C	SNOMED-FMA(body)	16.88	28.9	21.32
Vicuna-7B + BERT	C	SNOMED-FMA(body)	54.82	5.95	10.74
GPT-3.5 + Ada	C	SNOMED-NCIT(neoplas)	45.17	52.79	48.68
Falcon-7B + Ada	C	SNOMED-NCIT(neoplas)	42.73	52.42	47.08
Falcon-7B + BERT	C	SNOMED-NCIT(neoplas)	72.73	28.47	40.92
LLaMA-2-7B + Ada	C	SNOMED-NCIT(neoplas)	42.71	52.42	47.07
LLaMA-2-7B + BERT	C	SNOMED-NCIT(neoplas)	72.73	28.47	40.92
MPT-7B + Ada	C	SNOMED-NCIT(neoplas)	42.71	52.42	47.07
MPT-7B + BERT	C	SNOMED-NCIT(neoplas)	72.73	28.47	40.92
Mamba-2.8B + Ada	C	SNOMED-NCIT(neoplas)	38.23	46.0	41.76
Mamba-2.8B + BERT	C	SNOMED-NCIT(neoplas)	71.13	24.29	36.21
Mistral-7B + Ada	C	SNOMED-NCIT(neoplas)	44.47	52.39	48.11
Mistral-7B + BERT	C	SNOMED-NCIT(neoplas)	73.44	28.5	41.06
Vicuna-7B + Ada	C	SNOMED-NCIT(neoplas)	43.1	52.6	47.38
Vicuna-7B + BERT	C	SNOMED-NCIT(neoplas)	72.76	28.44	40.9

Table 23. LLM models results — BIO-ML track – Rep is the representation type. Retriever model Top-k is set to 5. PART 2

Model	Rep	Task	Results		
			Prec	Rec	F1
GPT-3.5 + Ada	C	SNOMED-NCIT(pharm)	81.41	58.13	67.83
Falcon-7B + Ada	C	SNOMED-NCIT(pharm)	78.48	58.38	66.96
Falcon-7B + BERT	C	SNOMED-NCIT(pharm)	92.08	20.64	33.73
LLaMA-2-7B + Ada	C	SNOMED-NCIT(pharm)	78.48	58.38	66.96
LLaMA-2-7B + BERT	C	SNOMED-NCIT(pharm)	92.08	20.64	33.73
MPT-7B + Ada	C	SNOMED-NCIT(pharm)	78.48	58.38	66.96
MPT-7B + BERT	C	SNOMED-NCIT(pharm)	92.08	20.64	33.73
Mamba-2.8B + Ada	C	SNOMED-NCIT(pharm)	73.41	50.99	60.18
Mamba-2.8B + BERT	C	SNOMED-NCIT(pharm)	91.66	17.04	28.74
Mistral-7B + Ada	C	SNOMED-NCIT(pharm)	79.52	58.26	67.25
Mistral-7B + BERT	C	SNOMED-NCIT(pharm)	92.07	20.61	33.68
Vicuna-7B + Ada	C	SNOMED-NCIT(pharm)	78.42	58.31	66.89
Vicuna-7B + BERT	C	SNOMED-NCIT(pharm)	92.15	20.64	33.73
GPT-3.5 + Ada	CC	NCIT-DOID(disease)	86.14	78.1	81.93
Falcon-7B + Ada	CC	NCIT-DOID(disease)	83.03	81.01	82.0
Falcon-7B + BERT	CC	NCIT-DOID(disease)	93.84	72.19	81.61
LLaMA-2-7B + Ada	CC	NCIT-DOID(disease)	82.95	80.88	81.9
LLaMA-2-7B + BERT	CC	NCIT-DOID(disease)	93.84	72.11	81.55
MPT-7B + Ada	CC	NCIT-DOID(disease)	83.06	81.07	82.05
MPT-7B + BERT	CC	NCIT-DOID(disease)	93.93	72.28	81.69
Mamba-2.8B + Ada	CC	NCIT-DOID(disease)	73.78	67.63	70.57
Mamba-2.8B + BERT	CC	NCIT-DOID(disease)	90.77	60.84	72.85
Mistral-7B + Ada	CC	NCIT-DOID(disease)	81.89	65.43	72.74
Mistral-7B + BERT	CC	NCIT-DOID(disease)	94.44	60.5	73.75
Vicuna-7B + Ada	CC	NCIT-DOID(disease)	81.98	79.3	80.62
Vicuna-7B + BERT	CC	NCIT-DOID(disease)	93.53	70.29	80.26
GPT-3.5 + Ada	CC	OMIM-ORDO(disease)	71.8	57.97	64.15
Falcon-7B + Ada	CC	OMIM-ORDO(disease)	69.05	58.51	63.34
Falcon-7B + BERT	CC	OMIM-ORDO(disease)	88.68	43.38	58.26
LLaMA-2-7B + Ada	CC	OMIM-ORDO(disease)	68.94	58.45	63.26
LLaMA-2-7B + BERT	CC	OMIM-ORDO(disease)	88.56	43.27	58.13
MPT-7B + Ada	CC	OMIM-ORDO(disease)	69.06	58.53	63.36
MPT-7B + BERT	CC	OMIM-ORDO(disease)	88.68	43.38	58.26
Mamba-2.8B + Ada	CC	OMIM-ORDO(disease)	63.06	51.01	56.4
Mamba-2.8B + BERT	CC	OMIM-ORDO(disease)	86.6	37.68	52.51
Mistral-7B + Ada	CC	OMIM-ORDO(disease)	69.39	43.43	53.42
Mistral-7B + BERT	CC	OMIM-ORDO(disease)	83.74	33.92	48.28
Vicuna-7B + Ada	CC	OMIM-ORDO(disease)	66.44	57.08	61.41
Vicuna-7B + BERT	CC	OMIM-ORDO(disease)	86.71	41.92	56.52
GPT-3.5 + Ada	CC	SNOMED-FMA(body)	17.56	29.56	22.03
Falcon-7B + Ada	CC	SNOMED-FMA(body)	16.84	28.82	21.26
Falcon-7B + BERT	CC	SNOMED-FMA(body)	54.7	5.94	10.72
LLaMA-2-7B + Ada	CC	SNOMED-FMA(body)	16.92	28.97	21.37
LLaMA-2-7B + BERT	CC	SNOMED-FMA(body)	54.81	5.97	10.76
MPT-7B + Ada	CC	SNOMED-FMA(body)	16.94	29.01	21.39
MPT-7B + BERT	CC	SNOMED-FMA(body)	54.81	5.97	10.76
Mamba-2.8B + Ada	CC	SNOMED-FMA(body)	16.71	28.96	21.19
Mamba-2.8B + BERT	CC	SNOMED-FMA(body)	54.24	5.55	10.08
Mistral-7B + Ada	CC	SNOMED-FMA(body)	13.1	4.42	6.61
Mistral-7B + BERT	CC	SNOMED-FMA(body)	42.31	0.76	1.49
Vicuna-7B + Ada	CC	SNOMED-FMA(body)	16.19	27.78	20.46
Vicuna-7B + BERT	CC	SNOMED-FMA(body)	53.88	5.55	10.07

Table 24. LLM models results — BIO-ML track – Rep is the representation type. Retriever model Top-k is set to 5. PART 3

Model	Rep	Task	Results		
			Prec	Rec	F1
GPT-3.5 + Ada	CC	SNOMED-NCIT(neoplas)	45.77	52.52	48.92
Falcon-7B + Ada	CC	SNOMED-NCIT(neoplas)	42.69	52.39	47.04
Falcon-7B + BERT	CC	SNOMED-NCIT(neoplas)	72.73	28.47	40.92
LLaMA-2-7B + Ada	CC	SNOMED-NCIT(neoplas)	42.68	52.39	47.04
LLaMA-2-7B + BERT	CC	SNOMED-NCIT(neoplas)	72.72	28.44	40.89
MPT-7B + Ada	CC	SNOMED-NCIT(neoplas)	42.64	52.31	46.98
MPT-7B + BERT	CC	SNOMED-NCIT(neoplas)	72.73	28.47	40.92
Mamba-2.8B + Ada	CC	SNOMED-NCIT(neoplas)	37.91	45.4	41.32
Mamba-2.8B + BERT	CC	SNOMED-NCIT(neoplas)	71.23	24.08	35.99
Mistral-7B + Ada	CC	SNOMED-NCIT(neoplas)	48.2	36.88	41.79
Mistral-7B + BERT	CC	SNOMED-NCIT(neoplas)	74.91	22.13	34.17
Vicuna-7B + Ada	CC	SNOMED-NCIT(neoplas)	42.59	51.84	46.76
Vicuna-7B + BERT	CC	SNOMED-NCIT(neoplas)	72.67	28.31	40.75
GPT-3.5 + Ada	CC	SNOMED-NCIT(pharm)	81.85	58.19	68.02
Falcon-7B + Ada	CC	SNOMED-NCIT(pharm)	78.53	58.37	66.96
Falcon-7B + BERT	CC	SNOMED-NCIT(pharm)	92.15	20.63	33.71
LLaMA-2-7B + Ada	CC	SNOMED-NCIT(pharm)	78.7	58.06	66.82
LLaMA-2-7B + BERT	CC	SNOMED-NCIT(pharm)	92.13	20.58	33.64
MPT-7B + Ada	CC	SNOMED-NCIT(pharm)	78.55	58.38	66.98
MPT-7B + BERT	CC	SNOMED-NCIT(pharm)	92.08	20.63	33.7
Mamba-2.8B + Ada	CC	SNOMED-NCIT(pharm)	72.96	50.35	59.58
Mamba-2.8B + BERT	CC	SNOMED-NCIT(pharm)	92.06	17.18	28.96
Mistral-7B + Ada	CC	SNOMED-NCIT(pharm)	85.19	56.21	67.73
Mistral-7B + BERT	CC	SNOMED-NCIT(pharm)	94.98	20.23	33.36
Vicuna-7B + Ada	CC	SNOMED-NCIT(pharm)	78.71	57.83	66.67
Vicuna-7B + BERT	CC	SNOMED-NCIT(pharm)	92.39	20.51	33.56
GPT-3.5 + Ada	CP	NCIT-DOID(disease)	85.52	72.24	78.32
Falcon-7B + Ada	CP	NCIT-DOID(disease)	83.04	80.99	82.0
Falcon-7B + BERT	CP	NCIT-DOID(disease)	93.89	72.17	81.61
LLaMA-2-7B + Ada	CP	NCIT-DOID(disease)	83.07	81.07	82.06
LLaMA-2-7B + BERT	CP	NCIT-DOID(disease)	93.9	72.28	81.68
MPT-7B + Ada	CP	NCIT-DOID(disease)	83.07	81.07	82.06
MPT-7B + BERT	CP	NCIT-DOID(disease)	93.9	72.28	81.68
Mamba-2.8B + Ada	CP	NCIT-DOID(disease)	74.41	67.65	70.87
Mamba-2.8B + BERT	CP	NCIT-DOID(disease)	90.34	59.45	71.71
Mistral-7B + Ada	CP	NCIT-DOID(disease)	88.28	73.77	80.38
Mistral-7B + BERT	CP	NCIT-DOID(disease)	94.87	67.09	78.6
Vicuna-7B + Ada	CP	NCIT-DOID(disease)	82.04	78.94	80.46
Vicuna-7B + BERT	CP	NCIT-DOID(disease)	93.65	70.53	80.46
GPT-3.5 + Ada	CP	OMIM-ORDO(disease)	74.52	50.93	60.5
Falcon-7B + Ada	CP	OMIM-ORDO(disease)	69.18	58.64	63.48
Falcon-7B + BERT	CP	OMIM-ORDO(disease)	88.74	43.4	58.29
LLaMA-2-7B + Ada	CP	OMIM-ORDO(disease)	69.18	58.64	63.48
LLaMA-2-7B + BERT	CP	OMIM-ORDO(disease)	88.74	43.4	58.29
MPT-7B + Ada	CP	OMIM-ORDO(disease)	69.18	58.64	63.48
MPT-7B + BERT	CP	OMIM-ORDO(disease)	88.74	43.4	58.29
Mamba-2.8B + Ada	CP	OMIM-ORDO(disease)	62.93	51.06	56.38
Mamba-2.8B + BERT	CP	OMIM-ORDO(disease)	85.57	36.82	51.48
Mistral-7B + Ada	CP	OMIM-ORDO(disease)	74.41	51.89	61.15
Mistral-7B + BERT	CP	OMIM-ORDO(disease)	89.31	39.75	55.01
Vicuna-7B + Ada	CP	OMIM-ORDO(disease)	69.05	57.81	62.93
Vicuna-7B + BERT	CP	OMIM-ORDO(disease)	88.44	42.54	57.45

Table 25. LLM models results — BIO-ML track – Rep is the representation type. Retriever model Top-k is set to 5. PART 4

Model	Rep	Task	Results		
			Prec	Rec	F1
GPT-3.5 + Ada	CP	SNOMED-FMA(body)	21.13	32.61	25.64
Falcon-7B + Ada	CP	SNOMED-FMA(body)	16.94	29.0	21.39
Falcon-7B + BERT	CP	SNOMED-FMA(body)	54.67	5.97	10.76
LLaMA-2-7B + Ada	CP	SNOMED-FMA(body)	16.94	29.0	21.38
LLaMA-2-7B + BERT	CP	SNOMED-FMA(body)	54.6	5.97	10.76
MPT-7B + Ada	CP	SNOMED-FMA(body)	16.93	29.0	21.38
MPT-7B + BERT	CP	SNOMED-FMA(body)	54.67	5.97	10.76
Mamba-2.8B + BERT	CP	SNOMED-FMA(body)	52.25	5.28	9.59
Mistral-7B + Ada	CP	SNOMED-FMA(body)	24.13	24.5	24.32
Mistral-7B + BERT	CP	SNOMED-FMA(body)	57.24	4.58	8.47
Vicuna-7B + Ada	CP	SNOMED-FMA(body)	16.91	28.94	21.35
Vicuna-7B + BERT	CP	SNOMED-FMA(body)	54.68	5.95	10.74
GPT-3.5 + Ada	CP	SNOMED-NCIT(neoplas)	46.96	52.26	49.47
Falcon-7B + Ada	CP	SNOMED-NCIT(neoplas)	42.71	52.42	47.07
Falcon-7B + BERT	CP	SNOMED-NCIT(neoplas)	72.73	28.47	40.92
LLaMA-2-7B + Ada	CP	SNOMED-NCIT(neoplas)	42.71	52.42	47.07
LLaMA-2-7B + BERT	CP	SNOMED-NCIT(neoplas)	72.73	28.47	40.92
MPT-7B + Ada	CP	SNOMED-NCIT(neoplas)	42.71	52.42	47.07
MPT-7B + BERT	CP	SNOMED-NCIT(neoplas)	72.73	28.47	40.92
Mamba-2.8B + Ada	CP	SNOMED-NCIT(neoplas)	38.94	47.03	42.61
Mamba-2.8B + BERT	CP	SNOMED-NCIT(neoplas)	72.76	25.0	37.21
Mistral-7B + Ada	CP	SNOMED-NCIT(neoplas)	45.18	46.4	45.78
Mistral-7B + BERT	CP	SNOMED-NCIT(neoplas)	73.17	26.81	39.25
Vicuna-7B + Ada	CP	SNOMED-NCIT(neoplas)	42.69	52.16	46.95
Vicuna-7B + BERT	CP	SNOMED-NCIT(neoplas)	72.81	28.44	40.91
GPT-3.5 + Ada	CP	SNOMED-NCIT(pharm)	81.95	54.08	65.16
Falcon-7B + Ada	CP	SNOMED-NCIT(pharm)	78.48	58.38	66.96
Falcon-7B + BERT	CP	SNOMED-NCIT(pharm)	92.08	20.64	33.73
LLaMA-2-7B + Ada	CP	SNOMED-NCIT(pharm)	78.48	58.38	66.96
LLaMA-2-7B + BERT	CP	SNOMED-NCIT(pharm)	92.08	20.64	33.73
MPT-7B + Ada	CP	SNOMED-NCIT(pharm)	78.48	58.38	66.96
MPT-7B + BERT	CP	SNOMED-NCIT(pharm)	92.08	20.64	33.73
Mamba-2.8B + Ada	CP	SNOMED-NCIT(pharm)	72.3	49.16	58.53
Mamba-2.8B + BERT	CP	SNOMED-NCIT(pharm)	91.42	16.89	28.51
Mistral-7B + Ada	CP	SNOMED-NCIT(pharm)	81.61	54.9	65.64
Mistral-7B + BERT	CP	SNOMED-NCIT(pharm)	92.82	19.15	31.74
Vicuna-7B + Ada	CP	SNOMED-NCIT(pharm)	78.55	58.18	66.84
Vicuna-7B + BERT	CP	SNOMED-NCIT(pharm)	92.04	20.52	33.56

Table 26. LLM models results — MSE track – Rep is the representation type. Retriever model Top-k is set to 5. PART 1

Model	**Rep**	**Task**	**Results**		
			Prec	Rec	F1
GPT-3.5 + Ada	C	MI-EMMO	90.32	88.89	89.6
Falcon-7B + Ada	C	MI-EMMO	75.31	96.83	84.72
Falcon-7B + BERT	C	MI-EMMO	96.67	92.06	94.31
LLaMA-2-7B + Ada	C	MI-EMMO	75.31	96.83	84.72
LLaMA-2-7B + BERT	C	MI-EMMO	96.67	92.06	94.31
MPT-7B + Ada	C	MI-EMMO	75.31	96.83	84.72
MPT-7B + BERT	C	MI-EMMO	96.67	92.06	94.31
Mamba-2.8B + Ada	C	MI-EMMO	72.5	92.06	81.12
Mamba-2.8B + BERT	C	MI-EMMO	96.43	85.71	90.76
Mistral-7B + Ada	C	MI-EMMO	89.55	95.24	92.31
Mistral-7B + BERT	C	MI-EMMO	96.61	90.48	93.44
Vicuna-7B + Ada	C	MI-EMMO	73.33	87.3	79.71
Vicuna-7B + BERT	C	MI-EMMO	96.49	87.3	91.67
GPT-3.5 + Ada	C	MI-MatOnto	71.11	21.19	32.65
Falcon-7B + Ada	C	MI-MatOnto	57.94	20.53	30.32
Falcon-7B + BERT	C	MI-MatOnto	89.71	20.2	32.97
LLaMA-2-7B + Ada	C	MI-MatOnto	57.94	20.53	30.32
LLaMA-2-7B + BERT	C	MI-MatOnto	89.71	20.2	32.97
MPT-7B + Ada	C	MI-MatOnto	57.94	20.53	30.32
MPT-7B + BERT	C	MI-MatOnto	89.71	20.2	32.97
Mamba-2.8B + Ada	C	MI-MatOnto	48.0	15.89	23.88
Mamba-2.8B + BERT	C	MI-MatOnto	89.29	16.56	27.93
Mistral-7B + Ada	C	MI-MatOnto	67.39	20.53	31.47
Mistral-7B + BERT	C	MI-MatOnto	89.71	20.2	32.97
Vicuna-7B + Ada	C	MI-MatOnto	54.29	18.87	28.01
Vicuna-7B + BERT	C	MI-MatOnto	89.23	19.21	31.61

Table 27. LLM models results — MSE track – Rep is the representation type. Retriever model Top-k is set to 5. PART 2

Model	**Rep**	**Task**	**Results**		
			Prec	Rec	F1
GPT-3.5 + Ada	CC	MI-EMMO	90.62	92.06	91.34
Falcon-7B + Ada	CC	MI-EMMO	75.31	96.83	84.72
Falcon-7B + BERT	CC	MI-EMMO	96.67	92.06	94.31
LLaMA-2-7B + Ada	CC	MI-EMMO	75.31	96.83	84.72
LLaMA-2-7B + BERT	CC	MI-EMMO	96.67	92.06	94.31
MPT-7B + Ada	CC	MI-EMMO	75.31	96.83	84.72
MPT-7B + BERT	CC	MI-EMMO	96.67	92.06	94.31
Mamba-2.8B + Ada	CC	MI-EMMO	74.07	95.24	83.33
Mamba-2.8B + BERT	CC	MI-EMMO	98.25	88.89	93.33
Mistral-7B + Ada	CC	MI-EMMO	94.74	85.71	90.0
Mistral-7B + BERT	CC	MI-EMMO	100.0	84.13	91.38
Vicuna-7B + Ada	CC	MI-EMMO	74.63	79.37	76.92
Vicuna-7B + BERT	CC	MI-EMMO	98.08	80.95	88.7
GPT-3.5 + Ada	CC	MI-MatOnto	69.57	21.19	32.49
Falcon-7B + Ada	CC	MI-MatOnto	57.55	20.2	29.9
Falcon-7B + BERT	CC	MI-MatOnto	89.55	19.87	32.52
LLaMA-2-7B + Ada	CC	MI-MatOnto	57.55	20.2	29.9
LLaMA-2-7B + BERT	CC	MI-MatOnto	89.55	19.87	32.52
MPT-7B + Ada	CC	MI-MatOnto	57.94	20.53	30.32
MPT-7B + BERT	CC	MI-MatOnto	89.71	20.2	32.97
Mamba-2.8B + Ada	CC	MI-MatOnto	51.0	16.89	25.37
Mamba-2.8B + BERT	CC	MI-MatOnto	89.47	16.89	28.41
Mistral-7B + Ada	CC	MI-MatOnto	70.51	18.21	28.95
Mistral-7B + BERT	CC	MI-MatOnto	94.92	18.54	31.02
Vicuna-7B + Ada	CC	MI-MatOnto	59.41	19.87	29.78
Vicuna-7B + BERT	CC	MI-MatOnto	89.23	19.21	31.61

Table 28. LLM models results — MSE track – Rep is the representation type. Retriever model Top-k is set to 5. PART 3

Model	**Rep**	**Task**	**Results**		
			Prec	Rec	F1
GPT-3.5 + Ada	CP	MI-EMMO	92.06	92.06	92.06
Falcon-7B + Ada	CP	MI-EMMO	75.31	96.83	84.72
Falcon-7B + BERT	CP	MI-EMMO	96.67	92.06	94.31
LLaMA-2-7B + Ada	CP	MI-EMMO	75.31	96.83	84.72
LLaMA-2-7B + BERT	CP	MI-EMMO	96.67	92.06	94.31
MPT-7B + Ada	CP	MI-EMMO	75.31	96.83	84.72
MPT-7B + BERT	CP	MI-EMMO	96.67	92.06	94.31
Mamba-2.8B + Ada	CP	MI-EMMO	74.07	95.24	83.33
Mamba-2.8B + BERT	CP	MI-EMMO	96.61	90.48	93.44
Mistral-7B + Ada	CP	MI-EMMO	98.0	77.78	86.73
Mistral-7B + BERT	CP	MI-EMMO	100.0	79.37	88.5
Vicuna-7B + Ada	CP	MI-EMMO	72.58	71.43	72.0
Vicuna-7B + BERT	CP	MI-EMMO	100.0	74.6	85.45
GPT-3.5 + Ada	CP	MI-MatOnto	75.95	19.87	31.5
Falcon-7B + Ada	CP	MI-MatOnto	57.94	20.53	30.32
Falcon-7B + BERT	CP	MI-MatOnto	89.71	20.2	32.97
LLaMA-2-7B + Ada	CP	MI-MatOnto	57.94	20.53	30.32
LLaMA-2-7B + BERT	CP	MI-MatOnto	89.71	20.2	32.97
MPT-7B + Ada	CP	MI-MatOnto	57.94	20.53	30.32
MPT-7B + BERT	CP	MI-MatOnto	89.71	20.2	32.97
Mamba-2.8B + Ada	CP	MI-MatOnto	56.12	18.21	27.5
Mamba-2.8B + BERT	CP	MI-MatOnto	89.8	14.57	25.07
Mistral-7B + Ada	CP	MI-MatOnto	86.67	17.22	28.73
Mistral-7B + BERT	CP	MI-MatOnto	91.07	16.89	28.49
Vicuna-7B + Ada	CP	MI-MatOnto	55.21	17.55	26.63
Vicuna-7B + BERT	CP	MI-MatOnto	87.93	16.89	28.33