Name:	***Silvana Yacoub***	***Maher Mohsen***
ID	*20201091*	*20200415*
Course:	***Processing of Formal and Natural Languages***	

Assignment 3

- ## Comparison between results of Pre-trained Model and Scratch Trained Model:

Apple:

Apple			
Pre-Trained			
Similar Words		Opposite Words	
Word	Similarity	Word	Similarity
Snapple	0.674534	QB1	-0.15394
appleyard	0.663798	Î²O	-0.15541
apples	0.62634	80,7.	-0.16573
applespoti	0.60808	D-x	-0.16777
applet	0.604103	ÅœPK	-0.17206
Nipple	0.589773	EÅ•Si	-0.17626
applespot	0.588238	Zipfi	-0.17943
apple-lisa	0.567416	qO	-0.17959
Rypple	0.553778	ERHS	-0.18481
Apple	0.548461	Ew	-0.18548

Apple			
Scratch			
Similar Words		Opposite Words	
Word	Similarity	Word	Similarity
applebee	0.710636914	gs	0.384939
apply	0.690165222	customize	0.340331
scrapple	0.687615216	stepup	0.30737
ripple	0.63644886	steve	0.305724
pineapple	0.558565617	yesterday	0.290853
appt	0.515812337	eh	0.289882
triple	0.447143286	stencil	0.2839
approval	0.422564358	customer	0.281207
applicable	0.406020492	interested	0.277923
ample	0.374791175	course	0.276421

Banana:

Banana			
Pre-Trained			
Similar Words		Opposite Words	
Word	Similarity	Word	Similarity
manana	0.791431	Å•KE	-0.05593
Melt-Banana	0.756687	.b	-0.07534
Manana	0.73553	573	-0.08744
Tanana	0.703046	Ig3	-0.09338
Bealanana	0.695887	G77	-0.10044
Sanana	0.67239	SNiP	-0.10055
Rienana	0.663097	eTI	-0.1033
Atsinanana	0.65791	555	-0.10678
Banana	0.651495	8206	-0.10811
znana	0.647766	p.3	-0.10812

Banana			
Scratch			
Similar Words		Opposite Words	
Word	Similarity	Word	Similarity
bananas	0.906352639	carpaccio	0.306893
ana	0.696177006	opi	0.290724
dana	0.565537035	round	0.281394
hana	0.551503062	easy	0.281071
alana	0.548280895	stud	0.27839
nirvana	0.523530662	informative	0.278011
tijuana	0.378118753	stock	0.276649
wcna	0.343515515	fitness	0.275239
divine	0.343358874	correspond	0.273449
banish	0.341878921	euro	0.272179

Carrot:

Carrot			
Pre-Trained			
Similar Words		Opposite Words	
Word	Similarity	Word	Similarity
parrot	0.789355	VÃµtu	0.191314
Barcarrotas	0.739765	M.T	0.175971
ParrotBebop	0.730022	FPK	0.147449
Puchowe	0.725651	iD`	0.132698
MDMA-d.Parrott	0.718509	JÃ¼	0.132466
garrote	0.708679	Bepco	0.132166
Parrott	0.701624	Ãœm	0.128625
Scaparrott	0.697271	.P	0.128012
carroures	0.693663	KÃ¶Ã¶	0.125833
Cottle	0.673952	bib	0.123073

Carrot			
Scratch			
Similar Words		Opposite Words	
Word	Similarity	Word	Similarity
carre	0.629027307	wawa	0.386813
arroz	0.505128086	hopefully	0.34141
carrier	0.500409842	bull	0.335568
rot	0.441484809	oar	0.310998
carte	0.422255099	subpar	0.303295
cart	0.408090562	raunchy	0.301087
arrow	0.372463584	wa	0.300551
carl	0.367175281	sushiand	0.292048
naked	0.361408621	familiar	0.291961
carman	0.35940969	ill	0.288118

Dog:

Dog			
Pre-Trained			
Similar Words		Opposite Words	
Word	Similarity	Word	Similarity
hot	0.713217	esialgsena	-0.246
food	0.557006	eesrindlikus	-0.24708
sugar	0.508893	osalisena	-0.24887
sucks	0.506942	ebatÃ¤ielikus	-0.24927
fucks	0.505417	kahebrigaadilisena	-0.25385
slutty	0.500833	autoriteedina	-0.25614
lazy	0.49334	fÃ¼Ã¼sikalisena	-0.25716
hairy	0.487935	aistitavana	-0.2622
drink	0.486259	temaatilisena	-0.26287
jumps	0.483448	kosmilises	-0.26601

Dog			
Scratch			
Similar Words		Opposite Words	
Word	Similarity	Word	Similarity
doggy	0.675956011	original	0.338421
og	0.62164855	removerweird	0.326707
doggie	0.576689899	parfait	0.326359
fog	0.532251537	coverage	0.318855
doc	0.434562325	hot	0.313637
log	0.432422191	beverage	0.312301
jog	0.42359069	slice	0.311053
uc	0.39166683	bullshit	0.308227
dock	0.390025705	origin	0.304616
dot	0.387280613	remodel	0.293834

Elephant:

Elephant			
Pre-Trained			
Similar Words		Opposite Words	
Word	Similarity	Word	Similarity
elephants	0.820782	EVC	-0.06539
Elephant	0.801807	Ä„gli	-0.08343
elephantopus	0.747856	QSL	-0.08672
Elephants	0.723123	Miy	-0.08763
ElÃ©phant	0.694809	eSLi	-0.08883
Aphantochroa	0.664167	ç□□å"¶éfi	-0.08953
decaying	0.642105	KMi	-0.09113
wood-decaying	0.641783	GÃ¤	-0.09307
towels	0.634109	Hitu	-0.09784
Elephantine	0.631994	ODUs	-0.0984

Elephant			
Scratch			
Similar Words		Opposite Words	
Word	Similarity	Word	Similarity
elegant	0.615853369	salsa	0.382636
thanno	0.460246772	mixed	0.352083
baron	0.440311491	solitary	0.341221
csk	0.437695503	salsayum	0.336823
stomache	0.408464015	marvelous	0.33362
thanks	0.406095564	dairy	0.314222
inashannon	0.389703959	miami	0.30932
stomach	0.38866064	salsas	0.309206
difficult	0.377113909	michigan	0.301037
stephanie	0.370117009	balsamic	0.290748

Flower:

Flower			
Pre-Trained			
Similar Words		Opposite Words	
Word	Similarity	Word	Similarity
wildflower	0.889169	loa-	-0.07092
Blower	0.839807	54-	-0.08224
Bellflower	0.830685	prg	-0.09646
ReyMayflower	0.830608	p.3	-0.09828
sunflower	0.82858	LOVL	-0.10724
blower	0.82366	lg3	-0.11129
Sunflower	0.815642	yelp	-0.11342
Wildflower	0.809431	TTTS	-0.11423
Flower	0.799598	0.72	-0.11565
Moonflower	0.798604	PHP4	-0.11636

Flower			
Scratch			
Similar Words		Opposite Words	
Word	Similarity	Word	Similarity
sunflower	0.826930702	declan	0.384936
lower	0.747795284	budino	0.316453
cauliflower	0.722260416	hibachi	0.311197
tower	0.611448944	hasselhoffe	0.30887
follower	0.561254919	deck	0.299655
flow	0.555751324	little	0.292172
power	0.542147219	similar	0.290821
shower	0.525895119	motivation	0.278213
flora	0.506107271	asopada	0.27195
floral	0.464147329	significantly	0.271902

Guitar:

Guitar			
Pre-Trained			
Similar Words		Opposite Words	
Word	Similarity	Word	Similarity
guitars	0.767422	0.72	-0.14034
drums	0.675521	lg3	-0.14463
Guitar	0.647187	ELs	-0.14468
Kitar	0.632662	7916	-0.15028
Sitar	0.627926	vÐ˜hem	-0.1531
goldrush	0.610903	p.3	-0.154
sitar	0.610868	FIEt	-0.15686
guitarra	0.60888	prg	-0.16095
vocal	0.602663	140st	-0.16851
swing	0.601596	SRKS	-0.17228

Guitar			
Scratch			
Similar Words		Opposite Words	
Word	Similarity	Word	Similarity
guitarist	0.578969955	nod	0.309296
tar	0.450099736	earl	0.296481
ohmygoodness	0.397105813	bibimbop	0.283452
tartar	0.395949066	budget	0.282493
oar	0.392498642	granola	0.270708
updo	0.377696425	budgetsavvy	0.268146
goodness	0.376236439	office	0.26621
restrictionsetc	0.370477498	ujs	0.26459
gooda	0.360880822	eyebrow	0.261106
sauvignon	0.359365702	budino	0.258516

House:

House			
Pre-Trained			
Similar Words		Opposite Words	
Word	Similarity	Word	Similarity
deep-house	0.757118	K-89	-0.21224
art-house	0.719953	L.-il	-0.21298
Icehouse	0.712557	PdLi	-0.21347
Dochouse	0.707282	35.1	-0.21573
warehouse	0.701123	LVÅ½	-0.21716
Epic-house	0.688885	Ì¼M	-0.22597
spouse	0.688839	XP31	-0.22791
Hothouse	0.688695	HÃ•	-0.23141
Henhouse	0.688074	3-lt	-0.23172
farmhouse	0.684496	9,3.	-0.23309

House			
Scratch			
Similar Words		Opposite Words	
Word	Similarity	Word	Similarity
madhouse	0.765157759	plan	0.321643
inhouse	0.703049719	anytime	0.31359
brewhouse	0.684699357	mile	0.304531
douse	0.673374355	allan	0.304326
townhouse	0.633041263	cilantro	0.287145
steakhouse	0.623654187	significantly	0.276911
housekeeper	0.621271193	lane	0.275219
rittenhouse	0.610117137	death	0.275113
smokehouse	0.602192521	bras	0.270667
mouse	0.566097081	miracle	0.269016

Internet:

Internet			
Pre-Trained			
Similar Words		Opposite Words	
Word	Similarity	Word	Similarity
m-internet	0.788116	Patxi	-0.14623
TVinternet	0.753582	Osc	-0.1556
Internet	0.75268	Agca	-0.15953
internett	0.751671	BayÃ³	-0.1674
Xinternet	0.739106	ç™½æ°å¸,	-0.16903
Koduinternet	0.707935	Ã“kÃ©r	-0.16939
nternet	0.688152	igl	-0.17256
interneta	0.668705	Ø¨Ø§Ù…ÙŠØ§Ù†	-0.17281
internetivÃµrk	0.658844	SÃ¤iz	-0.1742
internet.emt.ee	0.645397	58-3	-0.17445

Internet			
Scratch			
Similar Words		Opposite Words	
Word	Similarity	Word	Similarity
intern	0.877772152	laura	0.308778
internal	0.851165235	crabsjust	0.304799
international	0.723920047	lack	0.30075
interact	0.722835064	bartendress	0.295039
interior	0.711758614	margarhita	0.294098
interpret	0.651401877	salsas	0.287757
interview	0.644109905	las	0.283306
interpretation	0.56237179	crabmeat	0.280577
intense	0.516855419	money	0.276402
lantern	0.507934511	plateit	0.273227

Jungle:

Jungle			
Pre-Trained			
Similar Words		Opposite Words	
Word	Similarity	Word	Similarity
Bungle	0.699239	6116	-0.19979
jangle	0.651416	42a	-0.22003
Jangle	0.622974	7318	-0.22164
endAngle	0.613482	4766	-0.22678
Jungle	0.60771	8913	-0.22723
angle	0.599962	TPIK	-0.22984
mangle	0.598031	7431	-0.22988
Tangle	0.582189	37a.	-0.2305
Dangle	0.578572	280	-0.23067
triangle	0.576799	65609	-0.23311

Jungle			
Scratch			
Similar Words		Opposite Words	
Word	Similarity	Word	Similarity
jingle	0.703954756	benefit	0.308654
single	0.629478514	pad	0.284987
dingle	0.61215502	winebytheglass	0.284635
dangle	0.578898728	benedict	0.282613
eagle	0.488679796	turkey	0.279675
singletary	0.452142805	meat	0.278275
google	0.442208141	bonchon	0.277519
struggle	0.430143386	notice	0.275412
yuengle	0.429861993	wealth	0.267525
snuggle	0.391722232	weisswurst	0.267385

Conclusion:

The experiment underscores the effectiveness of pretrained word embeddings in capturing semantic information, highlighting their superiority over scratch-trained embeddings under similar training conditions.

- After training for 1500 epochs with a vector size of 100 and a window size of 10:
  - Pretrained word embeddings consistently outperformed scratch-trained embeddings.
- The results indicate that pretrained embeddings captured richer and more meaningful representations of words compared to scratch-trained embeddings.
- Leveraging pretrained embeddings offers advantages in terms of:
  - Learning robust semantic relationships between words.
  - Reducing training time and computational resources.
  - Facilitating faster model development and deployment.