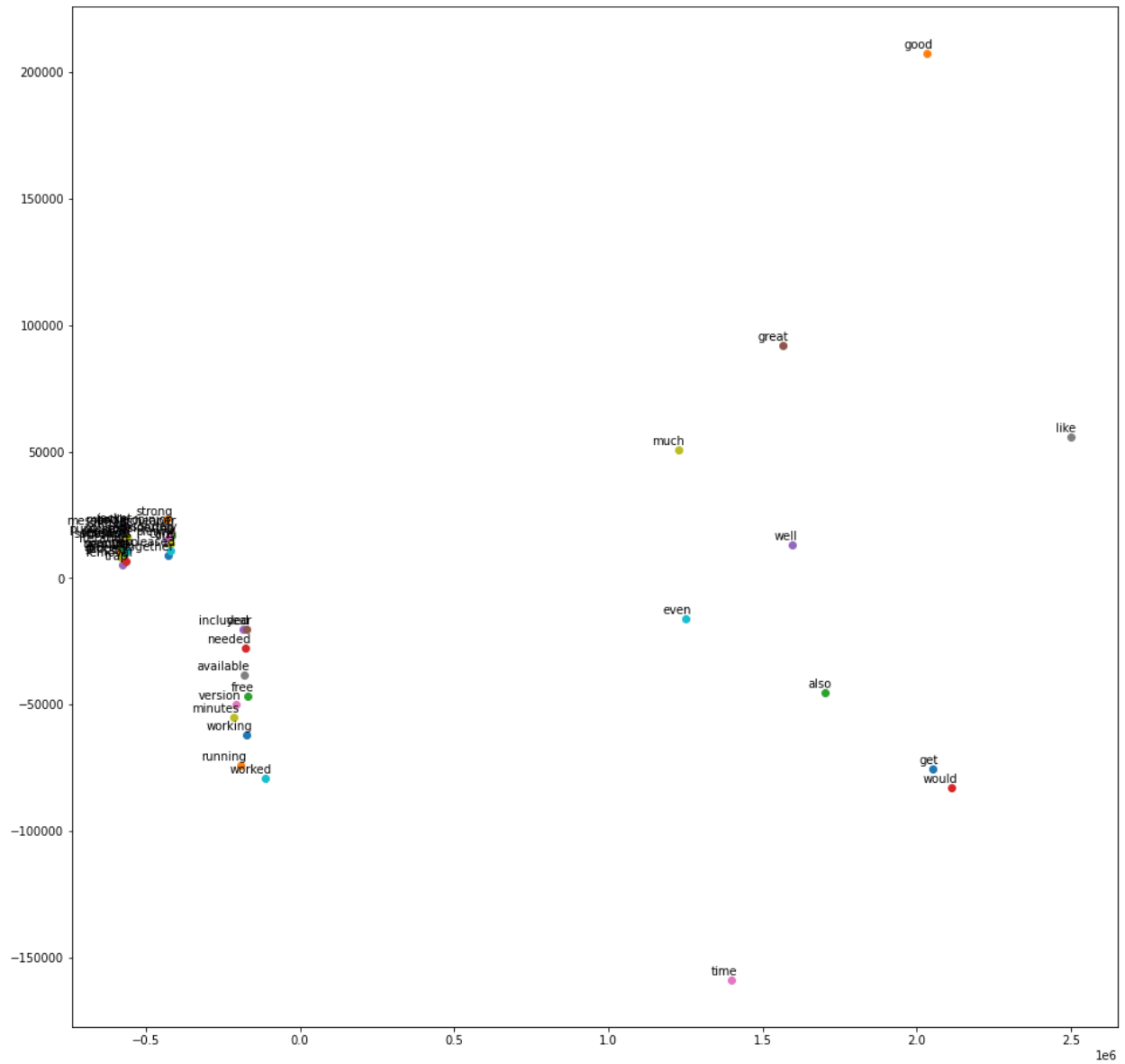# Advanced NLP Assignment 1

Mayank Goel, 2019114004

Question 1

Top closest words for Part (1)
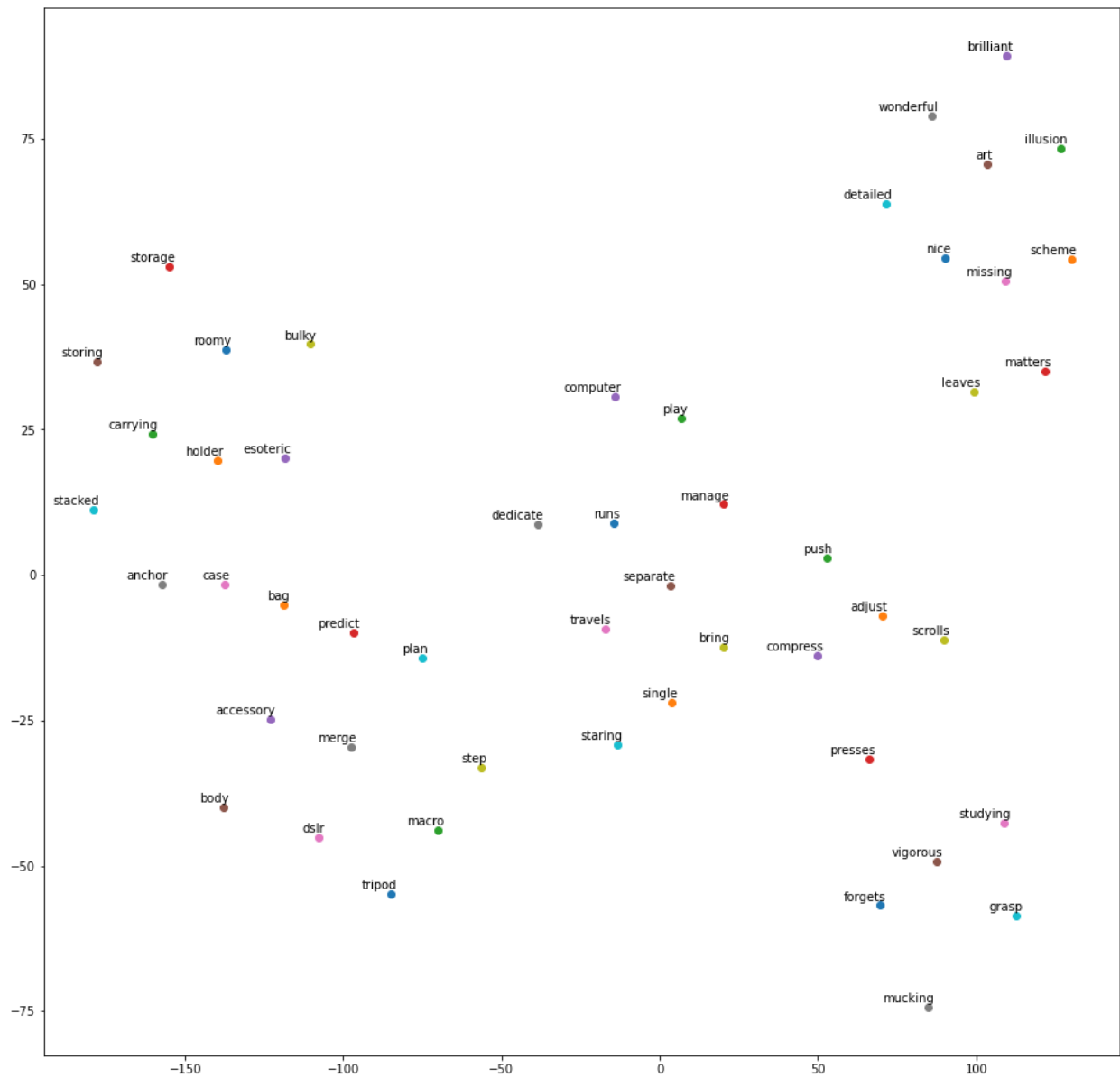
For camera:	For clean:	For depressed:	For run:	For wallet:
get	together	muscle	working	carbon
good	strong	springs	running	messenger
also	absolutely	iteration	free	genuine
would	care	grocery	needed	removal
well	plenty	trap	included	jacket
great	cheaper	sentence	year	snapped
time	considering	resistive	version	bottle
like	opinion	purposely	available	designs
much	tiny	tapped	minutes	squeeze
even	pleased	shifted	worked	gap

TSNE Plot for 50 word vectors

Top closest words for Part (2)

For camera:	For clean:	For depressed:	For run:	For wallet:
tripod	nice	forgets	runs	roomy
bag	scheme	adjust	single	holder
macro	illusion	push	play	carrying
predict	matters	presses	manage	storage
accessory	brilliant	compress	computer	esoteric
body	art	vigorous	separate	storing
dslr	missing	studying	travels	case
merge	wonderful	mucking	dedicate	anchor
step	leaves	scrolls	bring	bulky
plan	detailed	grasp	staring	stacked

TSNE Plot for 50 word vectors

Question 2

Co-occurence + SVD	CBOW	GloVe
get	tripod	cameras
good	bag	screen
also	macro	video
would	predict	screens
well	accessory	microphone
great	body	digital
time	dslr	images
like	merge	device
much	step	window
even	plan	sensor

We can see how the closest words in the co-occurence matrix are simply frequent words, and the result is largely non-sensical. In CBOW, we see a clear increase in accuracy, and the words seem meaningful, even if not as relevant as the one in the pretrained GloVe model.