Practical 9

Question 1:

All comments in the table are actual comments from some friends on Facebook. Raters are my unfortunate (fortunate) roommates.

Opinion	Rater 1	Rater 2	Rater3
I am far too dependent on my phone	neg	neg	neg
My phone is useful but not the best	neutral	neutral	pos
My phone is a bit shit - its second hand and the camera lens is cracked	neg	neg	neg

Cohen's kappa measures the inter-rater agreement which accounts for some agreement occurring by chance. It is calculated with the equation kappa = pr(a) – pr(e)/ 1 = pr(e) where pr(a) is the relative observed agreement among raters, and pr(e) is the hypothetical probability of chance agreement which is found by using the observed data to calculate the probabilities of each observer randomly saying each category. Alternatively, kappa can be found with the formula na – nc/n-nc where n is the number of subjects, na is number of agreements and nc = number of agreements due to chance. Complete agreement is shown with kappa = 1 and complete disagreement with kappa = 0.

		Rater 2				
		Positive	Negative	Neutral	Can't Say	Total
Rater 1	Positive	0	0	0	0	0
	Negative	0	2	0	0	2
	Neutral	0	1	0	0	1
	Can't Say	0	0	0	0	0
	Total	0	3	0	0	3
	Agreements	0	2	0	0	2
	Chance	.25	.25	.25	.25	1

So here between Rater 1 and Rater 2, kappa = na-nc / n – nc or 2-1/3-1 = ½ so kappa = 1.2 or .5

		Rater 3				
		Positive	Negative	Neutral	Can't Say	Total
Rater 1	Positive	0	0	0	0	0
	Negative	0	2	0	0	2
	Neutral	1	0	0	0	1
	Can't Say	0	0	0	0	0
	Total	1	2	0	0	3
	Agreement	0	2	0	0	2
	Chance	0.25	0.25	0.25	0.25	1

Here between Rater 1 and Rater 3, kappa = na-nc / n – nc or 2-1/3-1 = ½ so kappa = 1.2 or .5

		Rater 3				
		Positive	Negative	Neutral	Can't Say	Total
Rater 2	Positive	0	0	0	0	0
	Negative	1	2	0	0	3
	Neutral	0	0	0	0	0
	Can't Say	0	0	0	0	0
	Total	1	2	0	0	3
	Agreement	0	2	0	0	2
	Chance	0.25	0.25	0.25	0.25	1

So here between Rater 2 and Rater 3, kappa = na-nc/n-nc or 2 – 1/3 – 1 = ½ = .5 as well. In this case the inter-rater agreement is the same across the raters (.5). This is because they each rater had a different response to the same opinion (opinion 2) kappa is right in the middle of 1 and 0 so it is 50% reliable. Granted, this does not account for rater bias – the case that one judge might tend to prefer a certain category more than another.

We could also investigate inter-rater agreement using Pearson correlation. This correlation is a measure of the linear dependence or correlation between two variables. The coefficient will be somewhere in between 1 and -1 wherein 1 is a total positive and -1 is a total negative correlation; Any score close to either 1 or -1 is strong and get weaker as it approaches zero so a score of zero will mean that there is no correlation at all. To calculate Pearson's correlation, the agreement between each rater and the average of the remaining raters (in this case, the other two raters) must be measured. Then the average agreement over all raters must be measured.

Pearson's correlation coefficient (represented by rho) can be found by the following formula:

Rho x,y = cov(X,Y)/ standard deviation of x * standard deviation of y

If you plug in the formula of covariance into the above formula, you get the following:

expectation (x – mean of x)(y – mean of y) / standard deviation of x * standard deviation of y

To solve this, we would need to ascertain the standard deviation of both x and y, the means of both x and y, and the standard scores (which is calculated for x by the following formula: sample x – mean of x / standard deviation of x).

In summary, to find Pearson's Rho, the standard score, sample mean, and sample standard deviation must be calculated. Luckily, excel does this all in one function (PEARSONS array a, array b). However, first the text responses (the words positive, negative, neutral, can't say) must be changed to numerical values, in this case represented by 1s (if the rater picked this response) and 0s (if they didn't). However, calculating in this way revealed incomplete results, due to several division by zero errors. In fact, the only other result is 1, a complete positive correlation, found when the values in the given arrays were either identical, or where the array value for the rater was 1,0,1 and the Mean value was 1, .334, 1. In the cases where the division by zero error appeared, the standard deviation (the denominator of Pearson's correlation formula) was equal to zero. This is probably due to the several cells that do not have a value within them. This might be fixed if raters could give a score for each category from 1 to 10 (much like the example study with scores on the scale of 1 to 100). If raters are given an opportunity to give a variety of scores rather than just an absolute one or zero then perhaps more cells will have values and thus the standard deviation will be more than zero, thus fixing the division by zero error. "Can't Say" was cut as the raters never selected it. The resulting correlation table is as follows:

Opinion	Quality	Rater 1	Rater 2	Rater 3	Mean Rating
opinion 1	pos	0	0	0	0
opinion 2	pos	5	3	6	4.66666667
opinion 3	pos	0	0	0	0
opinion 1	neg	10	10	10	10
opinion 2	neg	5	7	4	5.33333333
opinion 3	neg	10	8	7	8.33333333
opinion 1	neutral	0	0	0	0
opinion 2	neutral	1	0	0	0.33333333
opinion 3	neutral	0	2	3	1.66666667
Positive	**Corr**	**Negative**	**Corr**	**Neutral**	**Corr**

R1xR2	1	R1xR2	0.75592895	R1xR2	-0.5
R1xR3	1	R1xR3	0.8660254	R1xR3	-0.5
R2xR3	1	R2xR3	0.98198051	R2xR3	1
R1xMR	1	R1xMR	0.93585673	R1xMR	-0.3273268
R2xMR	1	R2xMR	0.93812853	R2xMR	0.98198051
R3xMR	1	R3xMR	0.98666607	R3xMR	0.98198051

For the positive answers there is an almost complete positive correlation between all of the raters and all raters and the mean rating (all valued at 1). For the responses of negative, the correlations were all strong (strong/high correlation defined as greater than absolute value .5) and positively correlated (as one value increases, the other also increases). Rater three had the highest correlation with the mean rating and the lowest was the correlation between rater 1 and rater 2 (although the value of .756 is still categorized as a strong relationship). The negative correlations are found under the neutral category wherein Rater 1 and Rater 2 have a medium (absolute value of .3 to .5) negative correlation (a relationship wherein when the value of one variable increases, the value of the other decreases, and vice versa). The weakest correlation (weak defined as absolute value of less than .3) found was that of Rater 1 and the Mean Rating for the neutral quality with a negative correlation of -.327.

Question 2:

Three sentiment lists commonly used by researchers are:

- The General Inquirer (Harvard) *Positiv* and *Negativ* lists
- Hu and Liu's lexicon (University of Illinois Chicago)
- MPQA Opinion Corpus v3.0 (University of Pittsburgh)

<u>10 positive words from Hu and Liu's Lexicon</u>

Adventurous – Could be positive but also used occasionally as a polite demonization/judgement of another's lifestyle – in a sense can be used to describe those against the status quo (in terms of lifestyle choices, sex, etc.) Is also used as polite term for impulsive or non-conforming which may be negative E.g. referring to cuisine as "adventurous" – might have had lofty inspiration and dared to be different, but didn't quite make the mark. Kind of along the same lines as calling a piece of work "ambitious". Can be positive or negative depending on context. Also, as a side note, all adjectives can be countered with the addition of a negative word such as not. Not seeing this word in any context makes qualifying it difficult.

Gifted - always positive in that is positively quantifies the word that follows. Alone, this usually refers to high intelligence "She is quite gifted (i.e. smart). If I were to say "a gifted con-man" it is still positive in that I am quantifying and acknowledging a strong ability in a particular skill or area, regardless of how distasteful that skill/area may be.

Reforms – This is neither positive nor negative. It means change. Change can be good or bad depending upon context, person, and opinion.

Fascination – As above, this is neither positive nor negative, this merely quantifies interest. Interest in what is unknown, and it is the subject of interest that might better inform sentiment (positive/negative). Someone might have a morbid fascination with dead bugs, and that isn't positive, that is just weird. Again, context is key here.

Glamorous – generally a positive word although with adjectives that are generally positive, sarcasm might employ the use of these words for an opposite purpose. Context and tone is important here.

Temptingly – Alone this word implies that whatever verb is being modified (as this word is an adverb indicated by the presence of 'ly') is being done in an appealing way, but the positive or negative context is entirely dependent on the situation or verb its modifying (negative example: terms of seduction/ encouraging adultery, addition of drugs tempting a user, etc.

Problem-solver – usually good but in a context where emotional intelligence is required, using the word problem-solver to describe a person usually means they have little to no ability to connect on the emotional level. Again, context dependent

Worker - depends on the adjective before this word which modifies it. You can be a bad worker or lazy worker. Also can be a much more weighted word in terms of class discussions and political theory (think Marxism)

Hopeful - in theory a positive thing, but can be used sarcastically or to imply a misplaced hope which will never be realized (e.g. a hopeful actress)

Satisfied – in itself, this is always positive but the cause of the satisfaction can change the overall effect although the word itself is always positive

10 negative words from Hu and Liu's Lexicon

Crafty – can be a positive word denoting someone's ability to make arts and crafts well

Invalidate – always negative as in it does a negative active on the thing it is acting upon (usually a thought or feeling) but depending on the thought or feeling, this action could have a positive result (e.g. invalidating someone's feelings of worthlessness)

Loose – this is a neutral word, merely describing the state of grip, attachment, or adherence. A knot can be loose as can clothes, jewelry, or leaves of paper. These are not negative things.

Loopholes – Depending on a person's point of view these are either positive or negative – it entirely depends if a person benefits or suffers from them as a result. In context of a law it also entirely depends on a definition or positioning of moral compass.

Fumes – this is a neutral word and just describes a strong smell. Indeed strong smells often call up negative reactions but not all the time – for example, fumes from an incense burner.

Flakey – can be a positive word to describe a nice buttery well-baked pastry or crust

Debauchery – with the same definition can be used in a positive tone by people who want to go out and have a ridiculous session and not care about rules or adhering to any sort of responsibility. Often used with birthday party or friend holiday descriptions

Injury – always a negative word as in it always implies harm done to someone or something, but it is context dependent whether this harm is a good thing (e.g. harm done to a baddie so then the heroes win)

Junk – can be used to describe food or TV which is not inherently negative. It makes a judgement about the quality of the thing (e.g. junk food) but that does not make the thing itself any less appealing or desirable.

Dizzy – this is merely a state of being which can stem from negative occurrences (dehydration) or positive occurrences (having your first kiss)

10 Positive Words from Harvard Inquirer

Abide – completely context dependent. If the word "cannot" proceeds this word, it becomes a negative statement

Able – this merely means that something/someone can do something. That something remains to be seen and can completely change this from positive to negative (e.g. able to kill without guilt – yikes)

Affection – always a positive feeling towards a recipient but the presence of affection could be negative (e.g. Joe has a deep affection for his best friend's girlfriend)

Advocacy – a positive action for the recipient of the word but the overall valance depends on what you're advocating and also your own point of view. Free Health care? In my opinion, good - positive. In the opinions of ridiculous old rich white men in my country's senate? Bad - negative.

Affiliate - as above positive of the recipient of the action but depending on the recipient could be negative or positive. Affiliated with a school? Positive. Affiliated with war criminals? Negative.

Advisable – Always positive in itself when describing an entity. However as with other adjectives the inclusion of the word "not" beforehand can change everything.

Affluent – much like the word worker, usually a positive trait except in discussions of class or political theory

Afloat – always a good thing to that being modified by this adjective, however the overall context as in WHAT is afloat can change this from a positive to a negative thing.

Affirm – positive on the recipient of this action but overall valence depends on what is being affirmed. Belief in yourself? Positive. Belief that all government laws have no place in our lives? Probably negative.

Adorable – always a positive term but can be used with sarcasm or in a condescending manner where being adorable detracts (whether truly or merely perceived to detract) from ability. Example: calling a female colleague speaking at a meeting "adorable".

<u>10 Negative Words from Harvard Inquirer</u>

Abate – This can mean when something unpleasant becomes less intense, so it can actually be very positive

Abdicate – can be positive in the context of if a person wants to abdicate for a purpose (to marry an American divorcee) or maybe the people want a tyrant to abdicate and it happens.

Absurdity – absurdist comedy and absurdism are genera of literature and thus are neutral

Against – When used as a preposition completely neutral (I left the rake against the shed).

Alien – context dependent. In Sci-fi an alien is merely a type of character.

Alibi – Depends on the frame of reference. If you have an alibi, this is positive as you are not going to be arrested.

Addict – this has gone from a strictly negative word to a potential neutral word in slag being used to describe really liking something (e.g. book addict)

Admonition – This is neutral as it can just refer to a promise made (such as the final admonition a lady looking for a husband recites in the Disney movie "Mulan")

Antitrust – antitrust laws regulate businesses which, if you are a normal person should be regarded as a really positive thing.

Annihilate – means to completely eradicate and destroy so it is always negative for the recipient but depending on that recipient could be a positive thing e.g. annihilating world hunger. Can also be used as a synonym for completing winning/dominating a game which, depending on your team, can also be of positive valence.

Question 3:

To best quantify and judge the changes made on the given program's training, it is first necessary to establish a baseline. The changes made were to remove stop words, only look at words that were tagged as verbs, adjectives, adverbs, and interjections (because these words tend to show excitement or emotion) as these types of words tend to carry the most positive or negative meaning (a noun is not generally positive or negative in itself), and a combination of the two.

Stop words were removed by using the nltk.corpus stopwords "english" set and inserting the lines stops = set(stopwords.words("english")); posWords = [word for word in posWords if word not in stops] and ; negWords= [word for word in negWords if word not in stops].

Verbs, adjectives, adverbs, and interjections were found by writing a function that first found the part of speech by using nltk.pos_tag. Then the function would go through each word/POS tag tupple in the sentence and check the tag part of the tupple. Should the tag correspond to a tag of a verb, adjective, adverb, or interjection, then those words were added to a list which was then passed to the feature_select(words) part of the code. Only these words were added to the positive/negative word dictionary and thus the classifier was only trained with them. The code for the function is seen left.

```
def TagWords(sentence):
    a = list()
    try:
            tagged = nltk.pos_tag(sentence)
            for item in tagged:
                if item[1][0] == 'U': #interjection
                      a.append(item[0])
                elif item[1][0] =='V': #verbs
                    a.append(item[0])
                elif item [1][0] == 'R': #adverbs
                    a.append(item[0])
                elif item [1][0] == 'J': #adjectives
                    a.append(item[0])
    except Exception as e:
        print(str(e))
    print(a)
    return (a)
```

The four experimental cases are recorded in the below table. It is interesting to note that removing stopwords actually harmed all categories except for negative precision (although precision isn't always a good measure of a good classifier – you can hit the same entity all day due to high precision but that doesn't mean that entity is a correct hit). This could be because the stop-words users were in English and upon printing the sentences I am very sure that there were some Spanish sentences thrown in there. Additionally, stop words such as prepositions and sign-posting words like "but" and "therefore" add context, which, when attempting to determine the sentiment of a sentence are essential. Removing them changes the entire sentence and thus defeats the purpose of the experiment. This could also explain why removing both had negative effects in all categories. It seems that removing words that could at all provide context can change whether a sentence or even a word is deemed positive or negative. Further proof of the importance of word context can be seen in the explanations provided in question 2.

	Accuracy	Pos precision	Pos recall	Neg precision	Neg recall
No changes	.7734	.7881	.7479	.7602	.7990
Removing stopwords	.7648	.7658	.7629	.7638	.7667
Extracting only certain POS	.7303	.7452	.6999	.7171	.7607
Both	.706	.6942	.7359	.7190	.6759