NLP quiz 6, Name: 15 min , 10 marks

Q1.Match the following : [2 marks]

1.Homonym	a.Car is a ____ of vehicle.
2.Hyponym	b.Wheel is a ____ of car.
3.Meronym	c.Wood is a ____ of would.
4.Homephone	d.Bass (stringed instrument) , bass (fish) is instance of ____ .

1 - d
2 - a
3 - b
4 - c

Q2."I like backstreet boys , but not boys in the backstreet."

Calculate PMI(backstreet,boys) [2 marks]

Ans :

Total bigrams possible : 9 (=NB)
Total words :10 (=N)

Pr(backstreet,boys)= (count(backstreet,boys)/NB)=1/9=0.11
Pr(backstreet)=(count(backstreet)/N)=2/10=0.2
Pr(boys)=(count(boys)/N)=2/10=0.2

Plug into PMI formula,

PMI=$\log_2(0.11/(0.2*0.2))=\log_2(2.75)$

Q3. What is the problem with basic path-based similarity? Give an example based on Fig 1. What should be done to overcome this problem? [2 marks = 0.5 + 1(for the example) + 0.5] (Answer in not more that 3 lines)

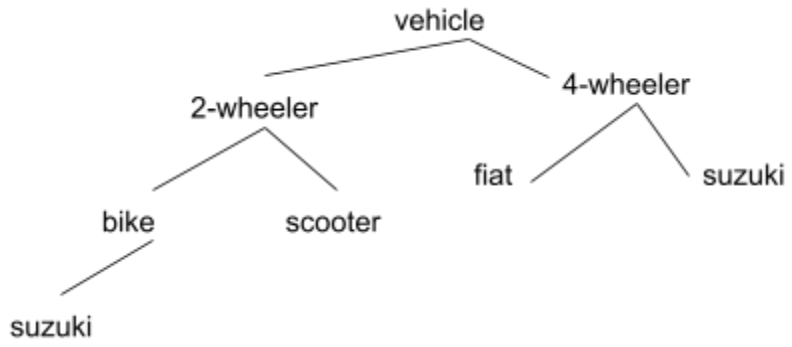
Ans :

Assumes each link represents a uniform distance although, nodes high in the hierarchy are very

abstract . E,g scooter seems to be closer to bike than to vehicle, but both have same similarity.

Soln :Represent the cost of each edge independently,such that words connected through abstract nodes are less similar.

Q4.Calculate wordsim(suzuki,vehicle) based on Fig 1 [4 marks]



(Fig 1)

1st sense of suzuki : sim($suzuki_1$,vehicle)= ¼
2nd sense of suzuki : sim($suzuki_2$,vehicle)=⅓

Thus sim(suzuki,vehicle)=max(sim1,sim2)=1/3