Natural Language Processing B.Tech VI Semester Section-3

u19cs080@coed.svnit.ac.in Switch account

Subjective Questions

Explain CKY(Cocke-Kasami-Younger) Algorithm.
Given a sentence “a gardener loves growing trees” and the phrase-structure grammar with probabilities in CNF(Chomsky normal form), generate all possible parse using CKY Algorithm for PCFG (Probabilistic Context Free Grammar).
Calculate the probabilities of all possible parse, and for the given sentence.

S	→ NP VP	[1.0]
VP	→ VBG NNS	[0.1]
VP	→ VBZ VP	[0.1]
VP	→ VBZ NP	[0.3]
NP	→ DT NN	[0.3]
NP	→ JJ NNS	[0.4]
DT	→ a	[0.3]
NN	→ gardener	[0.1]
VBZ	→ loves	[0.4]
VBG	→ growing	[0.5]
JJ	→ growing	[0.1]
NNS	→ trees	[0.34]

Add file

Calculate Minimum Edit Distance between words Occupation and Transition. Draw Table for calculating Levenshtein Distance and show all the steps.

Add file

Transition probability matrix

	Hot	Wet	Cold
Hot	0.6	0.3	0.1
Wet	0.4	0.4	0.2
Cold	0.1	0.4	0.5

Emission probability matrix

	Cotton	Nylon	Wool
Hot	0.8	0.5	0.05
Wet	0.15	0.4	0.2
Cold	0.05	0.1	0.75

o The above said matrix consists of emission probability values represented as $b_i(o_t)$. $b_i(o_t)$ is the probability of an observation o_t generated from a state b_i. For example, P(Nylon | Hot) = 0.5, P(Wool | Cold) = 0.75 etc.

$\pi = [\pi_1, \pi_2, \ldots, \pi_N]$ = set of prior probabilities = [0.6, 0.3, 0.1]. Here, the values refer to the probabilities P(Hot) = 0.6, P(Wet) = 0.3, and P(Cold) = 0.1

Consider P(Hot | <S>) = 0.4, P(Cold |<S>) = 0.3 and P(Wet |<S>) = 0.3

What would be the state sequence for observation (**cotton wool nylon cotton**) ? Do the calculations using Vertribi algorithm.

Add file

List out the phases/steps of Natural Language Processing. Explain the Syntactic Analysis Phase by considering the given simple context free grammar and sentence “take a cab from Station”. Create a parse tree using a top down approach with all possible steps.

S	→ NP VP	Aux	→ did | does
S	→ Aux NP VP	Det	→ the | a | this | that
S	→ VP	Prep	→ via | at | to | from | on
VP	→ Verb	Pronoun	→ me | he | she
NP	→ Pronoun	Proper-Noun	→ Airport | Station
Nominal	→ Noun	Noun	→ work | cab | guide
NP	→ Proper-Noun	Verb	→ involve | take | contain
VP	→ Verb NP		
VP	→ VP PP		
NP	→ Det Nominal		
Nominal	→ Nominal Noun		
Nominal	→ Nominal PP		
PP	→ Prep NP		

Add file

Consider a Corpus with 3 documents.

Document-1: In 2010, representation learning and deep neural network-style machine learning methods became widespread in natural language processing.

Document-2:That popularity was due partly to a flurry of results showing that such techniques can achieve state-of-the-art results in many natural language tasks, e.g., in language modeling and parsing.

Document-3:This is increasingly important in medicine and healthcare, where natural language processing helps analyze notes and text in electronic health records in other language, be inaccessible for study when seeking to improve care.

Using TF-IDF, Calculate rank order of all three documents for the word "**language**"

Add file

Back Submit Clear form