....

N.B.:
1) All questions are **COMPULSORY**.
2) Figures to the right indicate **FULL** marks.
3) Use of non-programmable **CALCULATOR** is allowed.
4) Assume suitable data if necessary.

Q.1 What is the smoothing for language model? Explain in detail. **[10]**

OR

Q.1 Explain the intrinsic evaluation metric (Perplexity) for N-gram model. **[10]**

Q.2 What is lemmatization and stemming? How is lemmatization done? Explain with example. **[10]**

OR

Q.2 What is disjunction grouping and precedence for pattern matching in regular expression? Explain with example. **[10]**

Q.3 Explain stochastic part of speech tagging. **[10]**

OR

Q.3 Define the following terms in morphology with example. **[10]**
   i) Morpheme   ii) Affix   iii) Prefixes   iv) Suffixes

Q.4 Explain probabilistic context free Grammar (CFG) in detail. **[10]**

OR

Q.4 Compare top-down and bottom-up parsing. **[10]**

Q.5 Write notes on following with example. **[10]**
   i) Homonymy   ii) Polysemy   iii) Hyponym   iv) Hepernym

OR

Q.5 Explain application of Natural Language Processing in text summarization. **[10]**

Q.6 Explain general architecture for text engineering (GATE). **[10]**

OR

Q.6 Write note on web ontology language (WOL). **[10]**

*   *   *   *