

CBCS SCHEME



Sixth Semester B.E/B.Tech. Degree Examination, Dec.2025/Jan.2026 Natural Language Processing

BAD613B

Max. Marks:100

Note: 1. Answer any FIVE full questions, choosing ONE full question from each module.
2. M : Marks , L: Bloom's level , C: Course outcomes.

| Module – 1 | | | M | L | C |
|-------------------|----|--|---|----|-----|
| 1 | a. | What is NLP? List and explain the application of NLP. | 6 | L2 | CO1 |
| | b. | Briefly explain various levels of natural language processing. | 7 | L2 | CO1 |
| | c. | List and explain the components of transformational grammar with an example. | 7 | L2 | CO1 |
| OR | | | | | |
| 2 | a. | What is NLP? Explain the challenges of NLP. | 5 | L2 | CO1 |
| | b. | Describe the Statistical Language Model (SLM) – n gram model. | 8 | L2 | CO1 |
| | c. | Explain the concept of Add-one smoothing and Good-Turing smoothing. Calculate the smoothed count for 5, by considering that the number of n-grams that occur 4 times is 25,108 and the number of n-grams that occur 5 times is 20,542. | 7 | L3 | CO1 |
| Module – 2 | | | | | |
| 3 | a. | Define a finite automation that accepts the following language $(aa)^*(bb)^*$. | 4 | L1 | CO2 |
| | b. | Compute the minimum edit distance between tutor and tumour. | 7 | L3 | CO2 |
| | c. | What is parsing? List and explain the different ways of word formation. | 9 | L2 | CO2 |
| OR | | | | | |
| 4 | a. | Write the regular expression for an email address. List and explain with example the use of square brackets. | 5 | L3 | CO2 |
| | b. | What is Morphological parsing? List and explain briefly the classes of morphemes. | 6 | L2 | CO2 |
| | c. | Write short notes on : i) Rule based Tagger ii) Stochastic Tagger iii) CYK Parser. | 9 | L2 | CO2 |

| BAD613B | | | | | | | | | | | | | | | | | | | | | | | |
|--|---------------------------------------|---|----|----|-----|-----|-----------|------------|-------------------|-----|---------------------------------------|------|----------------------------------|-------|---------------|------|---------------------------------|--------|-------------------------|-------|------|-----|-----|
| Module – 3 | | | | | | | | | | | | | | | | | | | | | | | |
| 5 | a. | Describe the Naïve Bayes algorithm. | 2 | L1 | CO3 | | | | | | | | | | | | | | | | | | |
| | b. | Derive the Naïve Bayes classifier equation. | 8 | L2 | CO3 | | | | | | | | | | | | | | | | | | |
| | c. | Test the following miniature training and test document from actual movie reviews using naïve Bayes with add-one smoothing. | 10 | L3 | CO3 | | | | | | | | | | | | | | | | | | |
| <table border="1" style="margin-left: auto; margin-right: auto;"> <thead> <tr> <th>cat</th> <th>Documents</th> </tr> </thead> <tbody> <tr> <td>Training –</td> <td>Just plain boring</td> </tr> <tr> <td>–</td> <td>Entirely predictable and lacks energy</td> </tr> <tr> <td>–</td> <td>No surprises and very few laughs</td> </tr> <tr> <td>+</td> <td>Very powerful</td> </tr> <tr> <td>+</td> <td>The most fun film of the summer</td> </tr> <tr> <td>Test ?</td> <td>Predictable with no fun</td> </tr> </tbody> </table> | | | | | | cat | Documents | Training – | Just plain boring | – | Entirely predictable and lacks energy | – | No surprises and very few laughs | + | Very powerful | + | The most fun film of the summer | Test ? | Predictable with no fun | | | | |
| cat | Documents | | | | | | | | | | | | | | | | | | | | | | |
| Training – | Just plain boring | | | | | | | | | | | | | | | | | | | | | | |
| – | Entirely predictable and lacks energy | | | | | | | | | | | | | | | | | | | | | | |
| – | No surprises and very few laughs | | | | | | | | | | | | | | | | | | | | | | |
| + | Very powerful | | | | | | | | | | | | | | | | | | | | | | |
| + | The most fun film of the summer | | | | | | | | | | | | | | | | | | | | | | |
| Test ? | Predictable with no fun | | | | | | | | | | | | | | | | | | | | | | |
| OR | | | | | | | | | | | | | | | | | | | | | | | |
| 6 | a. | Write an algorithm of Navie Bayes using add – 1 smoothing. | 7 | L3 | CO3 | | | | | | | | | | | | | | | | | | |
| | b. | With an example, explain binarization for the binary Naïve Bayes algorithm. | 6 | L2 | CO3 | | | | | | | | | | | | | | | | | | |
| | c. | Explain the Naïve Bayes as a language model. Illustrate the higher probability model for the following data set. | 7 | L3 | CO3 | | | | | | | | | | | | | | | | | | |
| <table border="1" style="margin-left: auto; margin-right: auto;"> <thead> <tr> <th>w</th> <th>P(w +)</th> <th>P(w -)</th> </tr> </thead> <tbody> <tr> <td>I</td> <td>0.1</td> <td>0.2</td> </tr> <tr> <td>Love</td> <td>0.1</td> <td>0.001</td> </tr> <tr> <td>This</td> <td>0.01</td> <td>0.01</td> </tr> <tr> <td>Fun</td> <td>0.05</td> <td>0.005</td> </tr> <tr> <td>film</td> <td>0.1</td> <td>0.1</td> </tr> </tbody> </table> | | | | | | w | P(w +) | P(w -) | I | 0.1 | 0.2 | Love | 0.1 | 0.001 | This | 0.01 | 0.01 | Fun | 0.05 | 0.005 | film | 0.1 | 0.1 |
| w | P(w +) | P(w -) | | | | | | | | | | | | | | | | | | | | | |
| I | 0.1 | 0.2 | | | | | | | | | | | | | | | | | | | | | |
| Love | 0.1 | 0.001 | | | | | | | | | | | | | | | | | | | | | |
| This | 0.01 | 0.01 | | | | | | | | | | | | | | | | | | | | | |
| Fun | 0.05 | 0.005 | | | | | | | | | | | | | | | | | | | | | |
| film | 0.1 | 0.1 | | | | | | | | | | | | | | | | | | | | | |
| CMRIT LIBRARY BANGALORE - 560 037 | | | | | | | | | | | | | | | | | | | | | | | |
| Module – 4 | | | | | | | | | | | | | | | | | | | | | | | |
| 7 | a. | With a neat diagram, explain the basic information retrieval process. | 6 | L2 | CO4 | | | | | | | | | | | | | | | | | | |
| | b. | List the different classical information retrieval models. And explain any one briefly. | 8 | L2 | CO4 | | | | | | | | | | | | | | | | | | |
| | c. | How does stemming affect the performance of an IR system? And what are the benefits of eliminating stop words. | 6 | L2 | CO4 | | | | | | | | | | | | | | | | | | |
| OR | | | | | | | | | | | | | | | | | | | | | | | |
| 8 | a. | List the different alternative models of IR. And explain any one briefly. | 8 | L2 | CO4 | | | | | | | | | | | | | | | | | | |
| | b. | Write short notes on : i) WORDNET ii) FrameNet iii)STEMMERS. | 12 | L2 | CO4 | | | | | | | | | | | | | | | | | | |

BAD613B

Module – 5

| | | | | | |
|----|----|--|----|----|-----|
| 9 | a. | Define Machine Translation. And explain the concept of: i) Word order typology ii) Lexical divergences iii) Morphological typology. | 10 | L2 | CO5 |
| | b. | With neat diagram, explain encode-decoder architecture. | 10 | L2 | CO5 |
| OR | | | | | |
| 10 | a. | List and explain the dimensions used to evaluate the transactions. And explain the Automatic Evaluation by Character Overlap (ChrF). | 10 | L2 | CO5 |
| | b. | Describe the bias and Ethical issues of machine translation. | 10 | L2 | CO5 |

CMRIT LIBRARY
BANGALORE - 560 037