

Mata Kuliah	:	PENGOLAHAN BAHASA ALAMI
Kelas	:	A
Pengampu	:	Dr. Retno Kusumaningrum, S.Si., M.Kom. Sukmawati Nur Endah, S.Si., M.Kom.
Jurusan	:	Informatika
Hari/Tanggal	:	Jumat, 8 April 2022
Jam/Ruang	:	15.20 – 17.00 (100 menit) / Online
Sifat Ujian	:	OPEN BOOK / COURSE MATERIAL

Jawablah pertanyaan-pertanyaan berikut pada lembar jawab yang telah disediakan!

(Angka pada awal setiap pertanyaan merupakan nilai dari soal yang bersangkutan)

1. (Skor : 20) Terapkan MED (*Minimum Edit Distance*) dari kata POLINOMIAL dan EKSPONENSIAL, tentukan berapa nilai MED dari kedua kata tersebut dan bagaimana alignment diantara kedua kata tersebut sehingga diperoleh skor MED yang sudah anda hitung sebelumnya!
2. (Skor : 10) Jelaskan apa yang dimaksud dengan *unigram* dan *bigram language modeling*!
3. (Nilai : 30) Jika diketahui sebuah *training data* sebagai berikut :

Id Doc	Class	Dokumen
1	Pos	Layan Bagus Makan Enak
2	Neg	Kamar Kotor Kamar Mandi Bau
3	Neg	Kamar Pengap Bau Kotor
4	Neg	Harga Mahal
5	Pos	Layan Bagus Makan Lengkap
6	Pos	Kamar Luas Layan Bagus

Dengan menggunakan *Multinomial Naive Bayes*, **tentukan dan jelaskan** kelas prediksi dari *testing document* berikut ini:

“Kamar Luas Artistik”

4. (Nilai : 20) Berikan contoh **kalimat opini/sentimen berbahasa Indonesia** dalam **bidang HOSPITALITY / PERHOTELAN** yang dapat digunakan untuk menunjukkan *issue* keterbatasan penggunaan *sentiment lexicon* seperti dijelaskan pada Tabel berikut ini!

Issue	Contoh
a. <i>Sentiment Word</i> dapat memiliki orientasi yang berbeda sesuai dengan domain pembicaraan, bisa bernilai positif atau negatif	
b. Sebuah kalimat mengandung <i>sentiment word</i> tetapi tidak mengekspresikan sentimen/opini apapun	
c. Kalimat sarkasme dengan atau tanpa <i>sentiment word</i> susah untuk diidentifikasi	
d. Kata tanpa <i>sentiment word</i> tetapi mengekspresikan suatu opini	

5. (Nilai : 20) Pada penerapan Latent Dirichlet Allocation sebagai metode *topic modeling* dikenal LDA sebagai *inference process* dan *generative process*. Jelaskan kedua proses tersebut dengan bahasa anda sendiri!