

		Universitas Negeri Surabaya Fakultas Matematika dan Ilmu Pengetahuan Alam Program Studi S1 Sains Data					Kode Dokumen											
RENCANA PEMBELAJARAN SEMESTER																		
MATA KULIAH (MK)		KODE	Rumpun MK		BOBOT (sks)		SEMESTER	Tgl Penyusunan										
Pemrosesan Teks		4920202030			T=3	P=0	ECTS=4.77	3 11 Agustus 2025										
OTORISASI		Pengembang RPS		Koordinator RMK		Koordinator Program Studi												
		Ulfa Siti Nuraini, M.Stat.		Dr. Atik Wintarti, M.Kom		YULIANI PUJI ASTUTI												
Model Pembelajaran	Project Based Learning																	
Capaian Pembelajaran (CP)	CPL-PRODI yang dibebankan pada MK																	
	CPL-3	Mengembangkan pemikiran logis, kritis, sistematis, dan kreatif dalam melakukan pekerjaan yang spesifik di bidang keahliannya serta sesuai dengan standar kompetensi kerja bidang yang bersangkutan																
	CPL-19	Mampu menerapkan pemikiran logis, kritis, sistematis, dan inovatif dalam mendesain, mengimplementasi dan mengevaluasi ilmu pengetahuan dan teknologi yang memperhatikan dan menerapkan nilai humaniora dalam bidang sains data.																
	CPL-22	Mampu merancang dan mengembangkan algoritma untuk berbagai keperluan seperti analisis big data, kecerdasan artifisial, basis data, penambangan data, statistika inferensial, desain dan analisis algoritma, dan data warehouse.																
	CPL-26	Menguasai konsep teoritis bidang pengetahuan sains data khusus dalam bidang pengetahuan tersebut secara mendalam, serta mampu memformulasikan penyelesaian masalah prosedural.																
	Capaian Pembelajaran Mata Kuliah (CPMK)																	
	CPMK - 1	Mampu merepresentasikan pengetahuan linguistik pada tingkat representasi morfologi, sintaksis serta semantik																
	CPMK - 2	Mampu melakukan penggalian data teks dari sumber digital dan mengolahnya menggunakan teknik pre-processing, ekstraksi fitur, dan similarity teks																
	CPMK - 3	Mampu melakukan pemodelan data teks dengan menggunakan klasifikasi dan kluster																
	CPMK - 4	Mampu merancang penyelesaian masalah pada data teks menggunakan pengolahan data teks yang terkait																
	Matrik CPL - CPMK																	
			CPMK	CPL-3	CPL-19	CPL-22	CPL-26											
		CPMK-1					✓											
		CPMK-2		✓														
	CPMK-3				✓													
	CPMK-4	✓																
Matrik CPMK pada Kemampuan akhir tiap tahapan belajar (Sub-CPMK)																		
		CPMK	Minggu Ke															
			1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
	CPMK-1		✓															
	CPMK-2			✓	✓	✓	✓	✓	✓									
	CPMK-3									✓	✓	✓						
	CPMK-4													✓	✓	✓	✓	✓
Deskripsi Singkat MK	Mata kuliah ini merupakan mata kuliah berbasis proyek yang mempelajari tentang teknik dasar untuk mengolah data teks. Pada mata kuliah akan diperkenalkan konsep morfologi bahasa, representasi teks, pre-processing, ekstraksi fitur hingga mendapatkan informasi seper similarity dan klasterisasi teks. Topik-topik yang dibahas meliputi: morfologi bahasa, represntasi string, regex, tokenisasi, text pre-processing, Bag of Words, TF-IDF, word similarity, word clustering, dan web scraping. Mahasiswa akan membuat proyek berkelompok untuk mengaplikasikan teori dan konsep pemrosesan teks pada permasalahan di bidang Sains Data.																	
Pustaka	Utama :																	
		1. Sarkar, Dipanjan, Text Analytics with Python A Practitioner's Guide to Natural Language Processing, Second Edition, Apress, 2019																
	Pendukung :																	

Deskripsi Singkat MK Mata kuliah ini merupakan mata kuliah berbasis proyek yang mempelajari tentang teknik dasar untuk mengolah data teks. Pada mata kuliah akan diperkenalkan konsep morfologi bahasa, representasi teks, pre-processing, ekstraksi fitur hingga mendapatkan informasi seperti similarity dan klusterisasi teks. Topik-topik yang dibahas meliputi: morfologi bahasa, representasi string, regex, tokenisasi, text pre-processing, Bag of Words, TF-IDF, word similarity, word clustering, dan web scraping. Mahasiswa akan membuat proyek berkelompok untuk mengaplikasikan teori dan konsep pemrosesan teks pada permasalahan di bidang Sains Data.

Pustaka

Utama :

1. Sarkar, Dipanjan, Text Analytics with Python A Practitioner's Guide to Natural Language Processing, Second Edition, Apress, 2019

Pendukung :

		1. Daniel Jurafsky & James H. Martin, Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition, 2nd Edition, Prentice Hall, 2008.					
Dosen Pengampu		Dr. Elly Matul Imah, M.Kom. Ike Fitriyaningsih, M.Si. Dr. Heri Purnawan, S.Si., M.Si. Riskiyana Dewi Intan Puspitasari, M.Kom. Ulfa Siti Nuraini, S.Stat., M.Stat. Dinda Galuh Guminta, M.Stat. Nurul Laili, S.Pd., M.Sc.					
Mg Ke-	Kemampuan akhir tiap tahapan belajar (Sub-CPMK)	Penilaian		Bantuan Pembelajaran, Metode Pembelajaran, Penugasan Mahasiswa, [Estimasi Waktu]		Materi Pembelajaran [Pustaka]	Bobot Penilaian (%)
		Indikator	Kriteria & Bentuk	Luring (offline)	Daring (online)		
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
1	Memahami konsep morfologi bahasa mulai dari fonem, morpheme, lexemes, syntax, dan contex	1.Menjelaskan konsep morfologi bahasa manusia 2.Menjelaskan konsep phoneme 3.Menjelaskan konsep morpheme dan lexeme 4.Menjelaskan konsep syntax 5.Menjelaskan konsep contex	Kriteria: Non-Tes Tugas Bentuk Penilaian : Aktifitas Partisipatif	Pembelajaran Kolaboratif (Ceramah, diskusi, dan tanya jawab) 3 X 50		Materi: Konsep Bahasa, Phoneme, Morpheme dan Lexeme, Syntax, Context Pustaka: <i>Daniel Jurafsky & James H. Martin, Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition, 2nd Edition, Prentice Hall, 2008.</i>	2%
2	Menjelaskan representasi string, operasi string, regex, dan visualisasi data teks pada pemrograman	1.Menjelaskan representasi string. 2.Menjelaskan operasi pada string 3. Menjelaskan teknik indexing dan slicing 4. Menjelaskan method pada string 5.Menjelaskan formatting pada string 6.Menjelaskan regex pada string	Kriteria: Non-Tes Tugas	Pembelajaran Kolaboratif (Ceramah, diskusi, dan tanya jawab) 3 X 50		Materi: Representasi string, Operasi string, Indexing dan slicing, String method, String formatting, String regex Pustaka: <i>Sarkar, Dipanjan, Text Analytics with Python A Practitioner's Guide to Natural Language Processing, Second Edition, Apress, 2019</i>	2%

3	Melakukan Web Scraping (HTTP dan API)	1. Mahasiswa mampu menjelaskan definisi dari web scraping 2. Mahasiswa mampu melakukan web scraping dengan HTTP 3. Mahasiswa mampu melakukan web scraping dengan API	Kriteria: Non-Tes Tugas Bentuk Penilaian : Aktifitas Partisipatif	Pembelajaran Kolaboratif (Ceramah, diskusi, dan tanya jawab) 3 X 50		Materi: Beautifulsoal, Website Layout, Crawling Pustaka: Sarkar, Dipanjan, Text Analytics with Python A Practitioner's Guide to Natural Language Processing, Second Edition, Apress, 2019	2%
4	Mahasiswa mampu melakukan teknik pre-prosesing data teks	1. Mahasiswa mampu melakukan teknik pre-processing data teks sederhana (tag HTML, stopwords, dll) 2. Mahasiswa mampu melakukan teknik pre-processing advance (parsing, case folding, stemming, dll)	Kriteria: Non-Tes Tugas	Pembelajaran Kolaboratif (Ceramah, diskusi, dan tanya jawab) 3 X 50		Materi: PreProcessing Pustaka: Sarkar, Dipanjan, Text Analytics with Python A Practitioner's Guide to Natural Language Processing, Second Edition, Apress, 2019	2%
5		1. Mahasiswa mampu melakukan ekstraksi fitur dengan one hot encoding 2. Mahasiswa mampu melakukan ekstraksi fitur dengan bag of words 3. Mahasiswa mampu melakukan ekstraksi fitur dengan TF-IDF	Kriteria: Non-Tes Tugas	Pembelajaran Kolaboratif (Ceramah, diskusi, dan tanya jawab) 3 X 50		Materi: Ekstraksi Fitur Pustaka: Sarkar, Dipanjan, Text Analytics with Python A Practitioner's Guide to Natural Language Processing, Second Edition, Apress, 2019	2%
6	Melakukan teknik ekstraksi fitur pada data teks	1. Mahasiswa mampu melakukan ekstraksi fitur dengan Word2Vec 2. Mahasiswa mampu melakukan ekstraksi fitur dengan Fast-Text 3. Mahasiswa mampu melakukan ekstraksi fitur dengan Glove	Kriteria: Tugas	Pembelajaran Kolaboratif (Ceramah, diskusi, dan tanya jawab) 3 X 50		Materi: Ekstraksi Fitur Pustaka: Sarkar, Dipanjan, Text Analytics with Python A Practitioner's Guide to Natural Language Processing, Second Edition, Apress, 2019	3%
7	Mahasiswa dapat melakukan teknik dalam similaritas dari data teks	1. Mahasiswa mampu melakukan teknik similarity text dengan term dan document 2. Mahasiswa mampu melakukan teknik similarity text dengan konsep jarak 3. Mahasiswa mampu melakukan teknik similarity text untuk recommender system	Kriteria: Non-Tes	Pembelajaran Kolaboratif (Ceramah, diskusi, dan tanya jawab) 3 X 50		Materi: Bag of Words, Bag of N-gram Words, TF-IDF Similarity Features Pustaka: Sarkar, Dipanjan, Text Analytics with Python A Practitioner's Guide to Natural Language Processing, Second Edition, Apress, 2019	3%

8	Ujian Tengah Semester	Ujian Tengah Semester	Kriteria: Non-Tes Bentuk Penilaian : Tes	Ujian Tengah Semester (UTS) 3 X 50		Materi: Semua materi sebelum UTS Pustaka: <i>Sarkar, Dipanjan, Text Analytics with Python A Practitioner's Guide to Natural Language Processing, Second Edition, Apress, 2019</i>	5%
9	Mahasiswa mampu melakukan teknik clustering text pada data teks	1.Mahasiswa mampu melakukan text clustering dengan hierarchical clustering 2.Mahasiswa mampu melakukan text clustering dengan non hierarchical clustering 3.Mahasiswa mampu melakukan text clustering dengan density clustering model	Kriteria: Non-Tes Bentuk Penilaian : Praktik / Unjuk Kerja	Pembelajaran Kolaboratif (Ceramah, diskusi, dan tanya jawab) 3 X 50		Materi: Text Clustering Pustaka: <i>Sarkar, Dipanjan, Text Analytics with Python A Practitioner's Guide to Natural Language Processing, Second Edition, Apress, 2019</i>	3%
10	Mahasiswa mampu mendapatkan topik utama dengan LDA	1.Mahasiswa mampu mengaplikasikan LDA pada data text 2.Mahasiswa mampu menjelaskan hasil LDA dari data teks	Kriteria: Non-Tes Bentuk Penilaian : Praktik / Unjuk Kerja	Pendekatan kolaboratif (diskusi dan ekspositori) Diskusi 3x50		Materi: Latent Dirichlet Allocation Pustaka: <i>Sarkar, Dipanjan, Text Analytics with Python A Practitioner's Guide to Natural Language Processing, Second Edition, Apress, 2019</i>	3%
11	Mahasiswa mampu menerapkan konsep NLP (pada summarization)	1.Mahasiswa mampu menjelaskan konsep NLP 2.Mahasiswa mampu melakukan summarization pada data teks	Kriteria: Non-Tes	Pendekatan kolaboratif (diskusi dan ekspositori) Diskusi 3x50		Materi: NLP Pustaka: <i>Sarkar, Dipanjan, Text Analytics with Python A Practitioner's Guide to Natural Language Processing, Second Edition, Apress, 2019</i>	3%
12	Mengimplementasikan teknik pemrosesan data pada permasalahan nyata	1.Menentukan topik permasalahan 2.Menyusun tim dan jadwal pengerjaan project 3.Membuat desain pemecahan masalah project	Kriteria: Tes Bentuk Penilaian : Penilaian Hasil Project / Penilaian Produk	Aktifitas Partisipasif, Praktik / Unjuk Kerja 3x50		Materi: Scrapping using HTTP, API, JSON Pustaka: <i>Sarkar, Dipanjan, Text Analytics with Python A Practitioner's Guide to Natural Language Processing, Second Edition, Apress, 2019</i>	10%

13	Mengimplementasikan teknik pemrosesan data pada permasalahan nyata	1. Presentasi kemajuan project 2. Review progress project 3. Membuat desain pemecahan masalah project	Kriteria: Presentasi Bentuk Penilaian : Penilaian Hasil Project / Penilaian Produk	Aktifitas Partisipasif, Praktik / Unjuk Kerja 3x50		Materi: Semua materi Pustaka: Sarkar, Dipanjan, <i>Text Analytics with Python A Practitioner's Guide to Natural Language Processing, Second Edition</i> , Apress, 2019	10%
14	Mengimplementasikan teknik pemrosesan data pada permasalahan nyata	1. Mengimplementasikan hasil project 2. Merealisasikan hasil project yang telah disetujui	Kriteria: Presentasi Bentuk Penilaian : Penilaian Hasil Project / Penilaian Produk	Aktifitas Partisipasif, Praktik / Unjuk Kerja 3x50		Materi: Semua materi Pustaka: Sarkar, Dipanjan, <i>Text Analytics with Python A Practitioner's Guide to Natural Language Processing, Second Edition</i> , Apress, 2019	10%
15	Mengimplementasikan teknik pemrosesan data pada permasalahan nyata	Presentasi kemajuan dan laporan pencapaian kerja kelompok	Kriteria: Presentasi Bentuk Penilaian : Penilaian Hasil Project / Penilaian Produk	Aktifitas Partisipasif, Praktik / Unjuk Kerja 3x50		Materi: Semua materi Pustaka: Sarkar, Dipanjan, <i>Text Analytics with Python A Practitioner's Guide to Natural Language Processing, Second Edition</i> , Apress, 2019	10%
16	Ujian Akhir Semester (UAS)	Presentasi dan tanya-jawab	Kriteria: Aktifitas Partisipasif, Penilaian Hasil Project / Penilaian Produk, Penilaian Portofolio Bentuk Penilaian : Aktifitas Partisipasif, Penilaian Hasil Project / Penilaian Produk, Penilaian Portofolio	Presentasi dan tanya jawab 3x50		Materi: Semua materi Pustaka: Sarkar, Dipanjan, <i>Text Analytics with Python A Practitioner's Guide to Natural Language Processing, Second Edition</i> , Apress, 2019	30%

Rekap Persentase Evaluasi : Project Based Learning

No	Evaluasi	Persentase
1.	Aktifitas Partisipasif	14%
2.	Penilaian Hasil Project / Penilaian Produk	50%
3.	Penilaian Portofolio	10%
4.	Praktik / Unjuk Kerja	6%
5.	Tes	5%
		85%

Catatan

1. **Capaian Pembelajaran Lulusan Prodi (CPL - Prodi)** adalah kemampuan yang dimiliki oleh setiap lulusan prodi yang merupakan internalisasi dari sikap, penguasaan pengetahuan dan ketrampilan sesuai dengan jenjang prodinya yang diperoleh melalui proses pembelajaran.
2. **CPL yang dibebankan pada mata kuliah** adalah beberapa capaian pembelajaran lulusan program studi (CPL-Prodi) yang digunakan untuk pembentukan/pengembangan sebuah mata kuliah yang terdiri dari aspek sikap, ketrampilan umum, ketrampilan

- husus dan pengetahuan.
3. **CP Mata kuliah (CPMK)** adalah kemampuan yang dijabarkan secara spesifik dari CPL yang dibebankan pada mata kuliah, dan bersifat spesifik terhadap bahan kajian atau materi pembelajaran mata kuliah tersebut.
 4. **Sub-CPMK Mata kuliah (Sub-CPMK)** adalah kemampuan yang dijabarkan secara spesifik dari CPMK yang dapat diukur atau diamati dan merupakan kemampuan akhir yang direncanakan pada tiap tahap pembelajaran, dan bersifat spesifik terhadap materi pembelajaran mata kuliah tersebut.
 5. **Indikator penilaian** kemampuan dalam proses maupun hasil belajar mahasiswa adalah pernyataan spesifik dan terukur yang mengidentifikasi kemampuan atau kinerja hasil belajar mahasiswa yang disertai bukti-bukti.
 6. **Kreteria Penilaian** adalah patokan yang digunakan sebagai ukuran atau tolok ukur ketercapaian pembelajaran dalam penilaian berdasarkan indikator-indikator yang telah ditetapkan. Kreteria penilaian merupakan pedoman bagi penilai agar penilaian konsisten dan tidak bias. Kreteria dapat berupa kuantitatif ataupun kualitatif.
 7. **Bentuk penilaian:** tes dan non-tes.
 8. **Bentuk pembelajaran:** Kuliah, Responsi, Tutorial, Seminar atau yang setara, Praktikum, Praktik Studio, Praktik Bengkel, Praktik Lapangan, Penelitian, Pengabdian Kepada Masyarakat dan/atau bentuk pembelajaran lain yang setara.
 9. **Metode Pembelajaran:** Small Group Discussion, Role-Play & Simulation, Discovery Learning, Self-Directed Learning, Cooperative Learning, Collaborative Learning, Contextual Learning, Project Based Learning, dan metode lainnya yg setara.
 10. **Materi Pembelajaran** adalah rincian atau uraian dari bahan kajian yg dapat disajikan dalam bentuk beberapa pokok dan sub-pokok bahasan.
 11. **Bobot penilaian** adalah prosentasi penilaian terhadap setiap pencapaian sub-CPMK yang besarnya proposional dengan tingkat kesulitan pencapaian sub-CPMK tsb., dan totalnya 100%.
 12. TM=Tatap Muka, PT=Penugasan terstruktur, BM=Belajar mandiri.

RPS ini telah divalidasi pada tanggal 3 Mei 2024

Koordinator Program Studi S1
Sains Data



YULIANI PUJI ASTUTI
NIDN 0031077804

UPM Program Studi S1 Sains
Data



NIDN 0021059403

File PDF ini digenerate pada tanggal 8 Desember 2025 Jam 21:40 menggunakan aplikasi RPS-OBE SiDia Unesa

