PARSHWANATH CHARITABLE TRUST'S



Subject Incharge: Prof. Sarala Mary Page No._

A.P. SHAH INSTITUTE OF TECHNOLOGY

Department of Computer Science and Engineering **Data Science**



Academic Year: 2023 2024 Subject : DAV THE SEVEN PRACTICE AREA OF TEXT ANALYTICS: Text mining can be divided into seven Practice areas, based on uneque characteristice of each area. Though distind, these arears are highly Interelated; a typical text mining project will require techniques from multiple areas. The seven practice areas are as follows: Bearch and Information Retrieval (IR): Storage and retrieval of text documents, including search engines and keyword search. 2) Document dustering: terms, emppele, Grouping and calegorizing paragraphs, or documents, using date mining dusting methode. 3) Document classification: Grouping and categorizing enippele, paragraph. or documents, using date mining classification methods, based on models models trained labeled examples. 4) Web Mining: Data and text mining on the Internet, with a specific

focus on the scale and interconnectedness of the web.

Department of CSE-Data Science | APSIT

PARSHWANATH CHARITABLE TRUST'S



A.P. SHAH INSTITUTE OF TECHNOLOGY

Department of Computer Science and Engineering
Data Science



Semester : Subject :		Academic Year: 20 - 20
5) Information Extraction	Activity .	STATE ROLL S
Identilication and	extraction	of orelevant facts
and relationships from process making structus	unetructur	ed text; the
process making structus	red dates,	from unstructured
and semistructured t	ext .	را در در در از این در از اواد
6) Natural Language Proces	sing: (NLP)	anghani anay
Low-level language	processing	and und eletandin
lasks (eq. lagging bot)	of speech), of the notation
synon ymously with con	pulational	linguistics.
Drust Extraction:	S. Berein	Law 2007pm.
Grouping of words and	phrases t	into semantically
emilar groupes	200	the love without
0 ,		bed at a while of