UNIT-5

CASE STUDIES.

Case studies in PA are real world examples that show how organizations use data, statistical models and ML techniques to predict future outcomes and make better decisions.

Survey Analysis: It is a text mining tool that reads open ended survey answers to find opinions, feelings etc.

1) Data Quality & Noise:

=> Informal language, errors, unstructured responses challenges processing

=> Essential preprocessing but difficult.

=> Grammatical mistakes common in open-ended responses.

2) Subjectivity & Sentiment:

=> Complex tones like +ve / -ve / neutral.

=> Algorithms miss full emotional range.

=> Mixed opinions Eg: 'great but pricey' : hard to categorize

3) Contextual Understanding:

=> Terms & meanings vary by product / context

=> Automated systems struggle

=> Feedback context affects interpretation.

Challenges:
- Informal language
- Subjectivity & Opinion Mining.
- Data privacy

4) Data Privacy:

=> Sensitive privacy personal info requires anonymization.

=> Security measures needed

=> confidential data in surveys demands protection

Question Answering: System that understands your natural language question and finds/gives the right answer from documents or knowledge bases.

1) Ambiguity & Polysemy:

=> words/ sentences have multiple meanings
=> User intent unclear
=> sentence structure adds confusion.

2) Semantic Understanding & Reasoning:

=> Beyond keywords: deep semantics, multi source answers.
=> complex extraction will be hard
=> Requires reasoning over the documents.

3) Domain Specificity:

=> Specialized terms (science/ legal) are not in general training.
=> Poor performances in niches.
=> Lacks domain terminology coverage

Challenges:
- Query Understanding
- Answer Extraction/ Formulation
- Real time processing.

4) Real-time Processing:

=> Instant responses needs efficient algorithms.
=> Powerful architectures are required
=> Handles the complex queries quickly.

5) Query Understanding:

=> Interpret natural language and intent accurately.
=> It determines expected answer type.

Challenges in Text Mining:

1) Data Quality & Noise:

=> Real-world text (surveys, social media) has spelling errors, abbrevations, informal language inconsistent formatting

=> Requires heavy preprocessing/ cleaning for accuray.

=> "Noise" significantly impacts mining results.

2) Language Ambiguity/Nuance:

=> Polysemy (multiple word meanings), idioms, cultural references confuse systems.

=> Needs advanced context-aware models.

=> Leads to the frequent misinterpretation.

3) Contextual Understanding:

=> Hard to grasp intent, Deep learning, implict relations.

=> Traditional methods fail on overall document meanning

=> Fails to capture implict relationships.

4) Scalability & Computational Costs:

=> Big data volumes demand high resources and scalable framework.

=> Limits are smaller organizations

=> Requires efficient big data processing

5) Lack of Domain-Specific Data:

=> General models poor in specialized domains (science, law) due to unique terminology.

=> Needs domain-adapted training

=> Poor performance on niche sentence structures.

Unit-1:

Computing Paradigms, Cloud Computing Fundamentals, Cloud Computing Architecture and management

Unit-2:

Cloud Deployment models, Cloud service M...
Technological Drivers for ...

Persuasion Analysis:

=> Persuasion Analysis in text mining involves identifying sentiment, and underlying themes in content such as marketing materials.

Identifying Persuasive Techniques:

=> Subtle the strategies (ethos, pathos, logos) beyond the keywords.
=> The deep semantic analysis needed.
=> Emotional / logic appeals hard to detect

Context-Aware models:

=> Depends on situation, audience, intent.
=> Highly variable effectiveness.
=> Cultural / target audience factors.

Lack of labeled Data:

=> Scarce datasets for training
=> manual annotation time-consuming, subjective.
=> Needs domain-specific creation.

Challenges:

1) Identifying Persuasive elements:

=> Subtle linguistic cues (emotional, appeals, credibility)
=> Distinguishes from facts / opinions.

2) Contextual Complexity:

=> Cultural / target audience dependent.
=> No universal models.