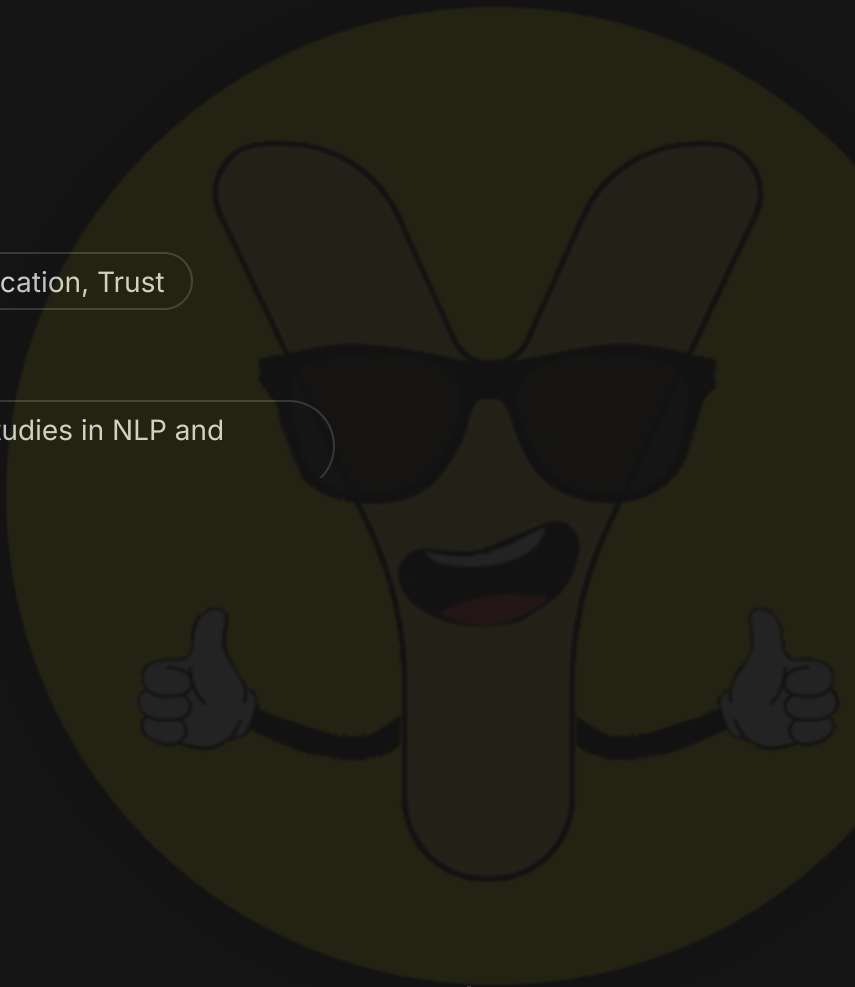
AKTU · 4TH YEAR · SEMESTER 7 · BCS701

Artificial Intelligence

Unit 5 — Applications of AI & Multi-Agent Systems

- Natural Language Processing
- Machine Translation
- Information Retrieval and Extraction
- Robotics: Perception, Planning, and Motion
- Speech Recognition
- Software Agents: Architecture, Communication, Trust
- Multi-agent Negotiation and Reputation
- Explainable AI (XAI) – Importance of interpretability, techniques for explaining black-box models, trust in AI, case studies in NLP and vision



DECODED BY

Chirag Jain

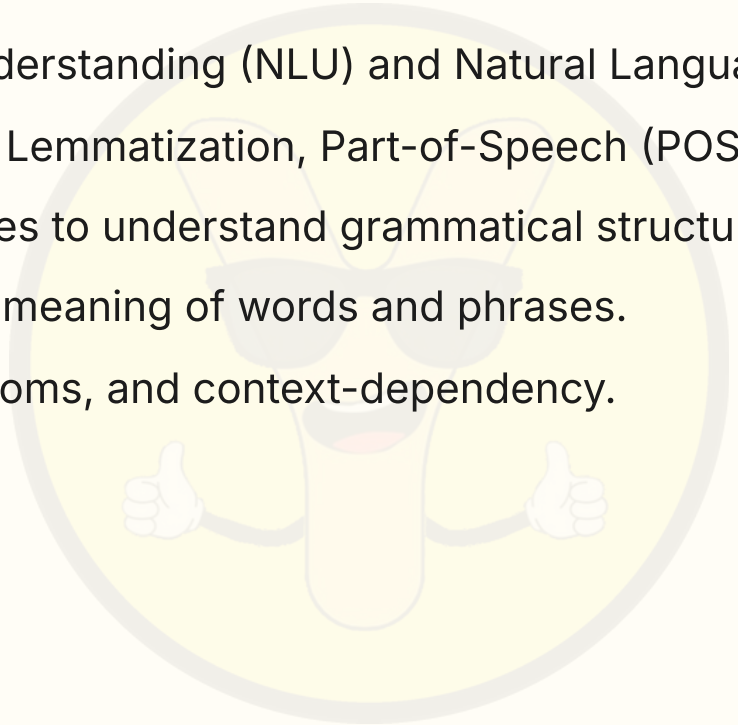
Founder, Ycotes · Gold Medalist in Artificial Intelligence, IIT Madras · taught 10,000+ students

theycotes.com

Natural Language Processing (NLP)

NLP is the field of AI that enables computers to understand, interpret, and manipulate human language.

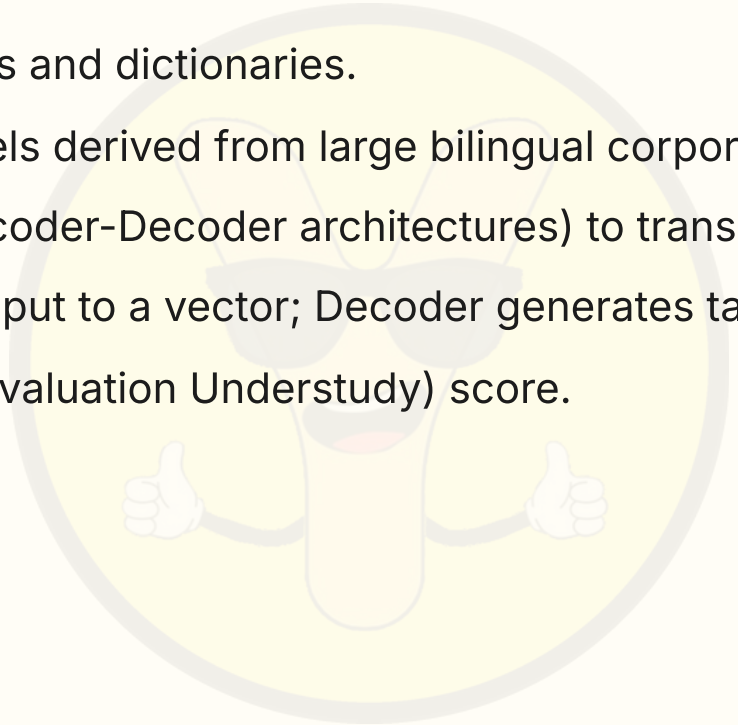
- Components: Natural Language Understanding (NLU) and Natural Language Generation (NLG).
- Key tasks: Tokenization, Stemming, Lemmatization, Part-of-Speech (POS) tagging.
- Syntactic analysis: Parsing sentences to understand grammatical structure.
- Semantic analysis: Determining the meaning of words and phrases.
- Challenges: Ambiguity, sarcasm, idioms, and context-dependency.



Machine Translation

The process of using software to translate text or speech from one natural language to another.

- Rule-based MT: Uses linguistic rules and dictionaries.
- Statistical MT: Uses statistical models derived from large bilingual corpora.
- Neural MT: Uses deep learning (Encoder-Decoder architectures) to translate sequences.
- Encoder-Decoder: Encoder maps input to a vector; Decoder generates target sequence.
- Evaluation metric: BLEU (Bilingual Evaluation Understudy) score.



Information Retrieval and Extraction

Retrieval focuses on finding relevant documents, while Extraction focuses on pulling structured data from unstructured text.

- Information Retrieval (IR): Searching for information in a large collection (e.g., Search Engines).
- IR Metrics: Precision = (Relevant Retrieved / Total Retrieved), Recall = (Relevant Retrieved / Total Relevant).
- Information Extraction (IE): Identifying entities (NER) and relationships.
- Named Entity Recognition (NER): Classifying entities into categories like Person, Organization, Location.
- Relation Extraction: Identifying semantic links between entities.

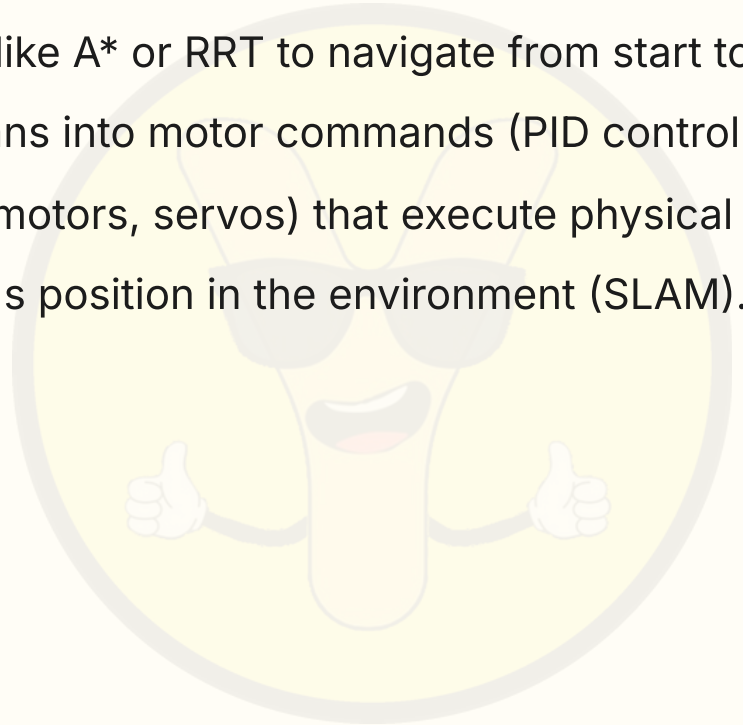
Robotics: Perception, Planning, and Motion



The standard robotic control loop showing information flow from environment to action.

Robotics: Core Components

- Perception: Interpreting sensor data (LIDAR, Cameras) to build a world model.
- Planning: Path planning algorithms like A* or RRT to navigate from start to goal.
- Motion Control: Converting path plans into motor commands (PID controllers).
- Actuators: Hardware components (motors, servos) that execute physical movement.
- Localization: Determining the robot's position in the environment (SLAM).



Speech Recognition

The process of converting spoken language into text.

- Acoustic Modeling: Maps audio signals to phonemes (basic units of sound).
- Language Modeling: Predicts the probability of word sequences.
- Hidden Markov Models (HMM): Traditionally used to model temporal speech sequences.
- Deep Learning: Current state-of-the-art uses RNNs, LSTMs, and Transformers.
- Challenges: Background noise, accents, and varying speech rates.

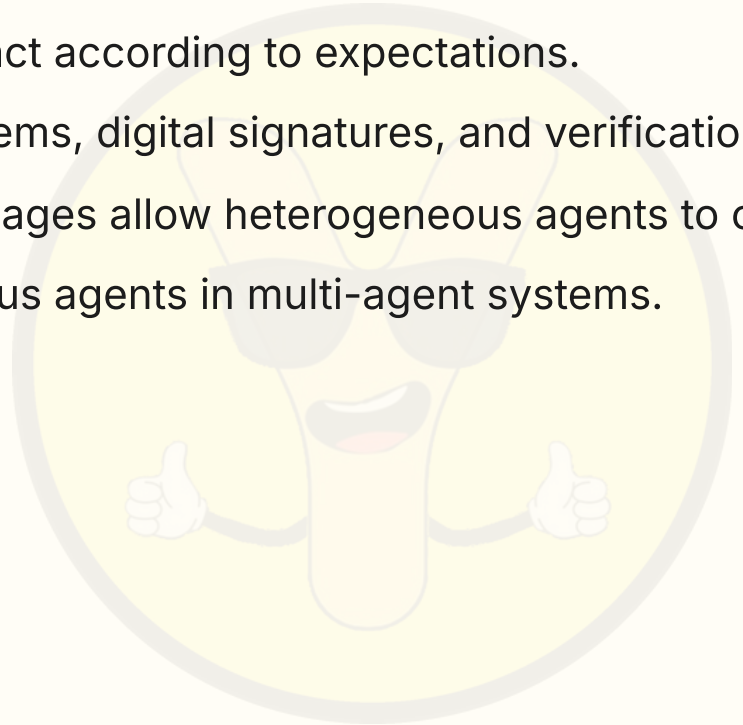
Software Agents: Architecture

An agent is an entity that perceives its environment and acts upon it.

- Simple Reflex Agents: Act only on current percepts (if-then rules).
- Model-based Reflex Agents: Maintain internal state to handle partial observability.
- Goal-based Agents: Act to achieve specific goals.
- Utility-based Agents: Act to maximize a performance measure (utility function).
- Learning Agents: Improve performance over time through experience.

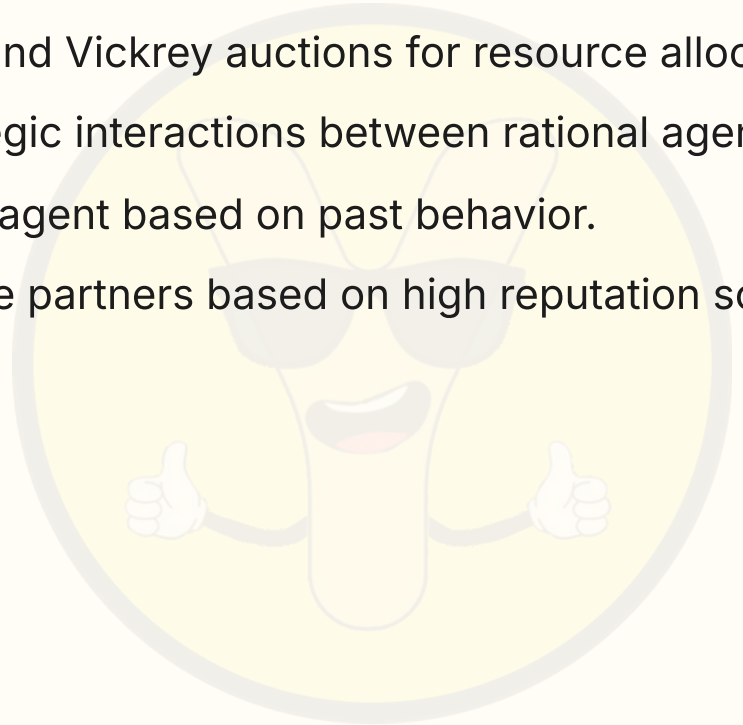
Agent Communication and Trust

- Communication: Agents exchange information using protocols like KQML or FIPA-ACL.
- Trust: The belief that an agent will act according to expectations.
- Trust mechanisms: Reputation systems, digital signatures, and verification protocols.
- Interoperability: Standardized languages allow heterogeneous agents to collaborate.
- Security: Protection against malicious agents in multi-agent systems.



Multi-agent Negotiation and Reputation

- Negotiation: The process where agents reach an agreement on a common goal.
- Auction Protocols: English, Dutch, and Vickrey auctions for resource allocation.
- Game Theory: Used to model strategic interactions between rational agents.
- Reputation: A score assigned to an agent based on past behavior.
- Trust-based models: Agents choose partners based on high reputation scores.

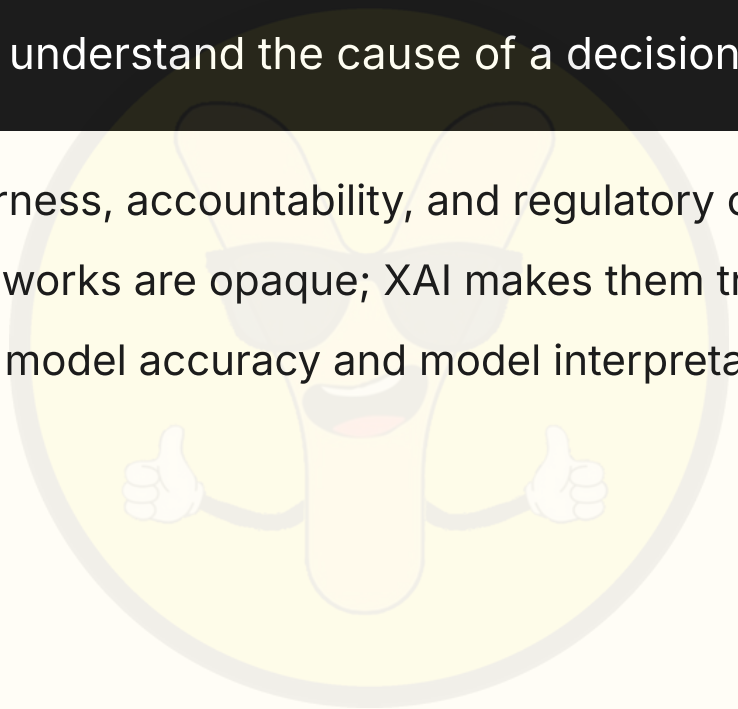


Explainable AI (XAI)

INTERPRETABILITY

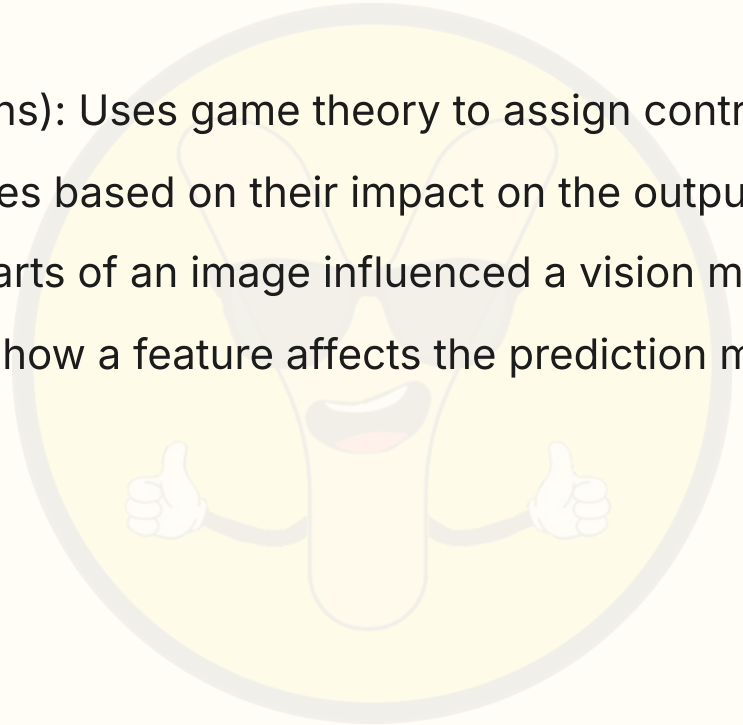
The degree to which a human can understand the cause of a decision made by an AI system.

- Importance: Essential for safety, fairness, accountability, and regulatory compliance.
- Black-box models: Deep Neural Networks are opaque; XAI makes them transparent.
- Trade-off: Often a conflict between model accuracy and model interpretability.



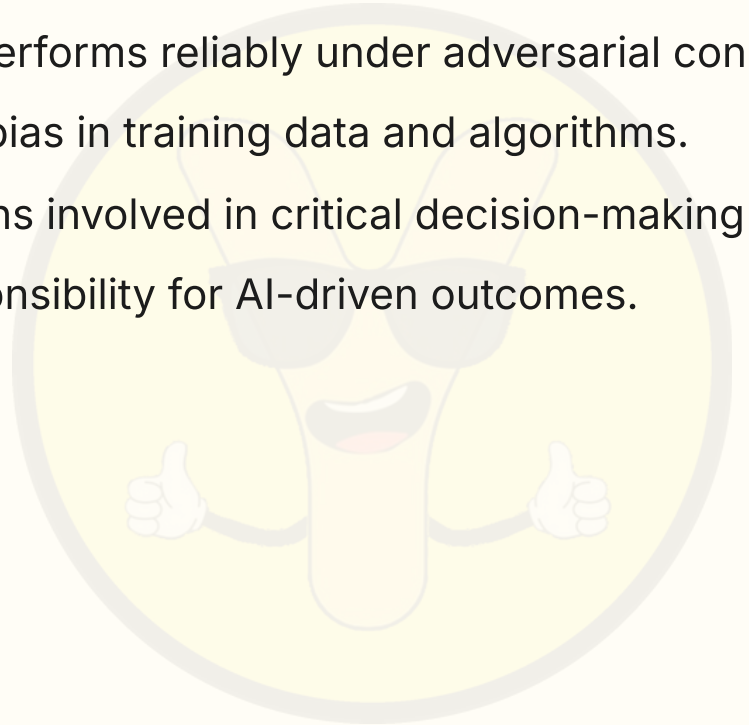
Techniques for Explaining Black-box Models

- LIME (Local Interpretable Model-agnostic Explanations): Approximates complex models with simple linear ones locally.
- SHAP (SHapley Additive exPlanations): Uses game theory to assign contribution values to features.
- Feature Importance: Ranking features based on their impact on the output.
- Saliency Maps: Visualizing which parts of an image influenced a vision model's prediction.
- Partial Dependence Plots: Showing how a feature affects the prediction marginalizing others.



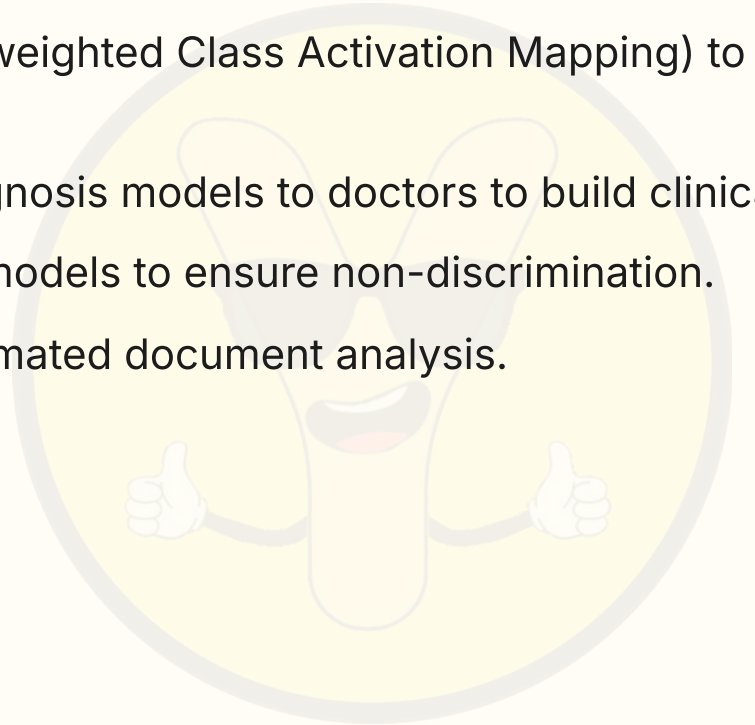
Trust in AI

- Transparency: Providing clear documentation of model training and logic.
- Robustness: Ensuring the system performs reliably under adversarial conditions.
- Fairness: Detecting and mitigating bias in training data and algorithms.
- Human-in-the-loop: Keeping humans involved in critical decision-making processes.
- Accountability: Defining clear responsibility for AI-driven outcomes.



Case Studies: NLP and Vision

- NLP Case: Using attention weights in Transformers to explain which words influenced a translation.
- Vision Case: Grad-CAM (Gradient-weighted Class Activation Mapping) to highlight image regions for classification.
- Healthcare: Explaining medical diagnosis models to doctors to build clinical trust.
- Finance: Explaining credit scoring models to ensure non-discrimination.
- Legal: Providing reasoning for automated document analysis.



Exam corner

- **Q1** Explain the architecture of a goal-based agent with a neat diagram.
- **Q2** Differentiate between Precision and Recall in Information Retrieval with formulas.
- **Q3** What is the role of the Encoder-Decoder architecture in Machine Translation?
- **Q4** Describe the importance of XAI in modern AI systems.
- **Q5** Explain the working of LIME or SHAP for black-box model interpretation.
- **Q6** Discuss the challenges involved in robotic perception and path planning.
- **Q7** How do reputation systems facilitate trust in multi-agent environments?

Unit 5 done. You are exam-ready.

Revise the Exam corner once more, then open Unit 1 on theycotes.com. Every AKTU subject, unit by unit, free. Notes by Chirag Jain, Ycotes.

