AGENDA

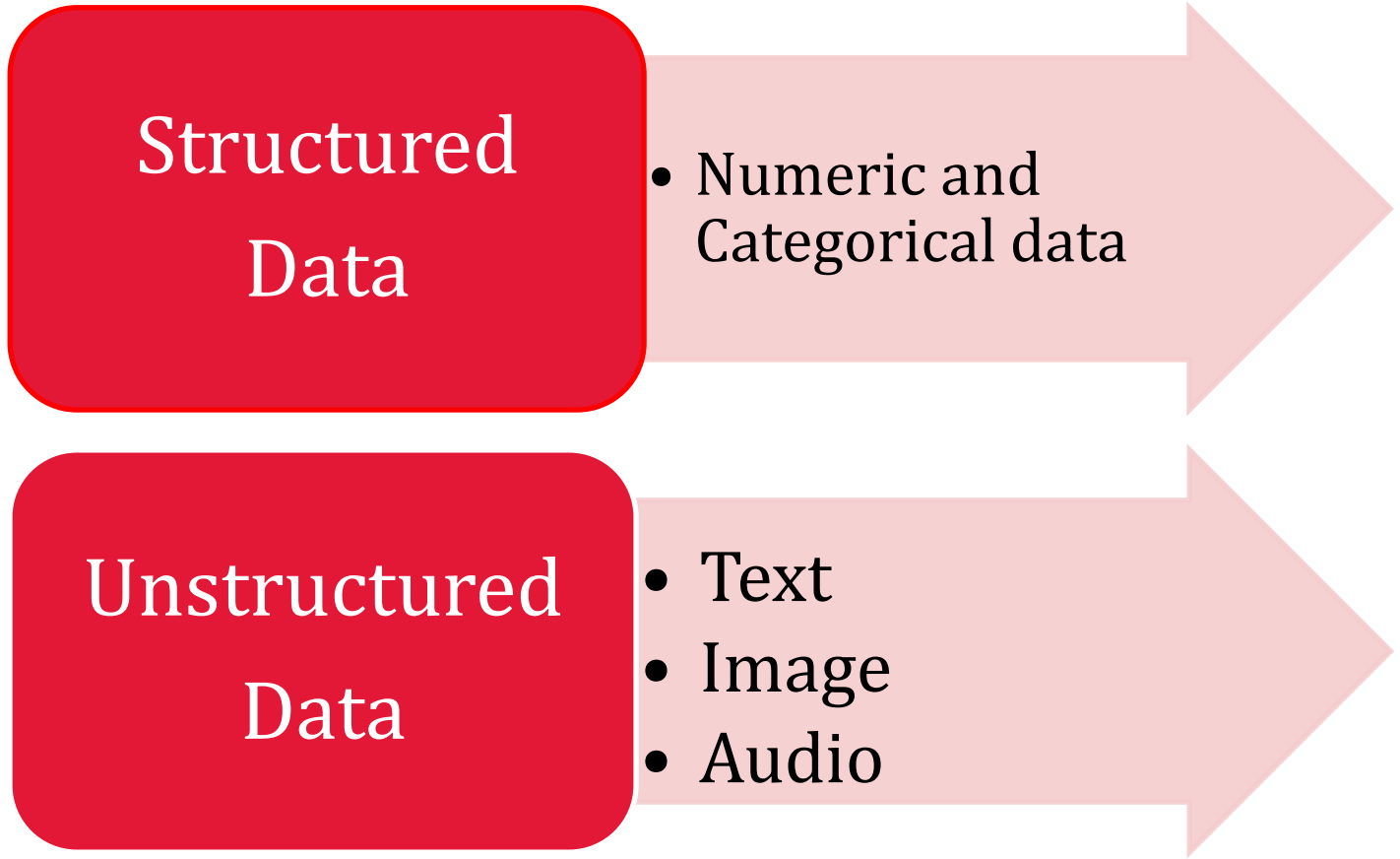
- Some business problem scenarios
- KNIME Solutions
- Categorization of problems
- Why KNIME?
- KNIME Demo - Install / Workbench / Workflow
- KNIME Learning resources
- KNIME Model Deployment & Post-Deployment
- How to identify opportunities in your accounts

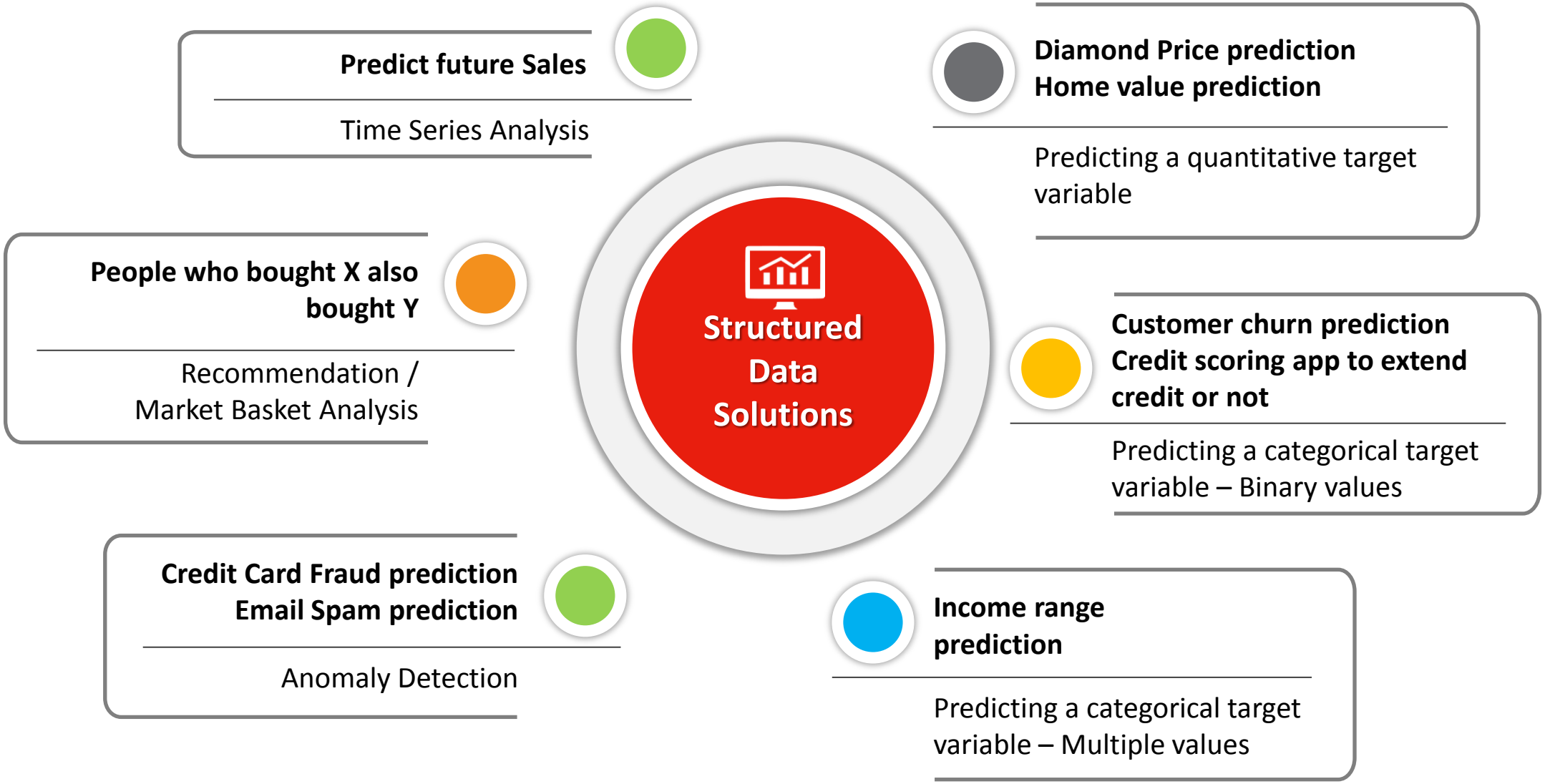
Some Business Problem Scenarios

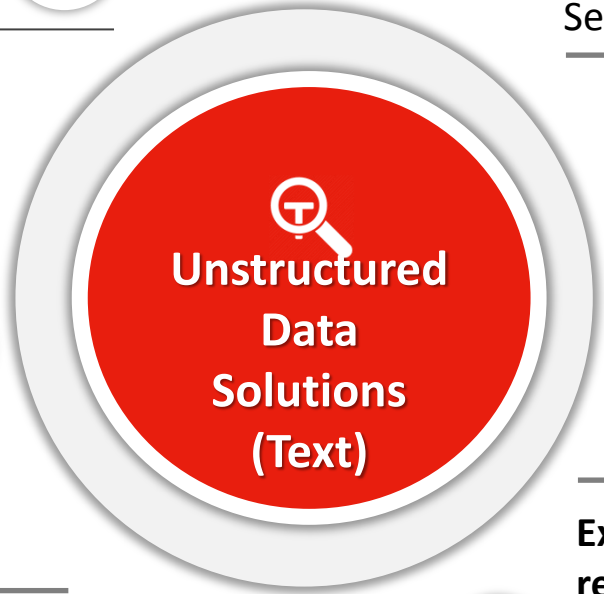
Available Input	Required Output
Diamond attributes – Cut, Carat, Clarity, Color	Estimated Price
Credit card transaction attributes – Amount, spending location, past spending average etc	Is the transaction likely to be fraudulent?
Customer relationship attributes – recent service feedback, spending pattern, complaints etc	Is the customer likely to leave?
Support incident Description	Category of the incident
Customer requests for account statements by email	Get account number and statement period from email body text
Product reviews in social media	Are the reviews positive or negative?
Images of products being inspected visually for abnormalities	Inspection pass or fail?
Images from a lab test showing presence of cells	Cell count

Missing business rules connecting input to output

Historical data (input and output) available







Advanced Query of Documents by meta data as well as content

Q&A

Sentiment Analysis of product review / employee feedback/movie review

Sentiment Analysis

Grouping of similar topic incidents based on description
Grouping News stories under different topics

Topic modelling

Given Incident description, assigning to correct category
Map User input to User Intent in Chatbots

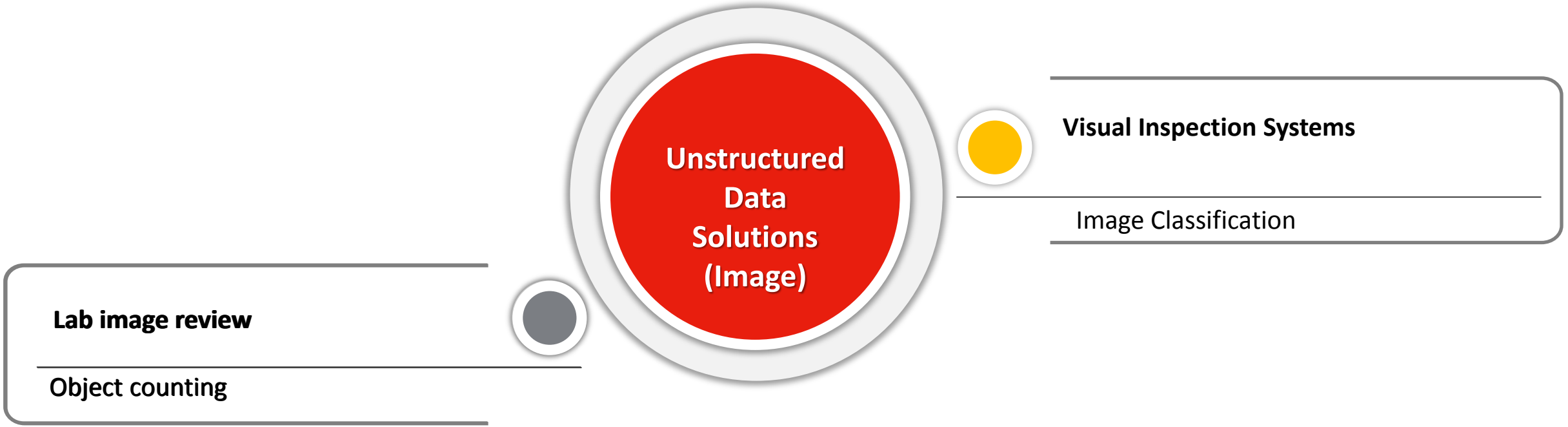
Document Classification, Chatbot Intent classification

Extract Names of PERSONS, ORGANIZATIONS, LOCATIONS, DATES, NUMBERS, MONEY, PERCENTAGE from any document

Named Entity Extraction using commercial APIs

Extract external auditor name from Annual reports
Extract customer a/c number and date period for a statement request through email

Custom Entity Extraction - Annotate training documents, train model, Entity extract using model



Why KNIME?

- Gartner Magic Quadrant
- Open Source
- No file size restrictions (Rapid Miner 10K)
- Responsive User Forum – Queries responded to within couple of days
- Comparatively easier to learn and get used to
- Rich functionality – Over 1500 nodes + Community & Partner extensions
- Integration with Python, R, Java, H2O
- Integration with Deep Learning Frameworks – Keras, DL4J

KNIME Demo

1. Install Video - https://www.youtube.com/watch?v=yeHblDxakLk
2. KNIME Work Bench Walkthru - https://www.youtube.com/watch?v=A32NoHC4Uf8
3. KNIME Learning with Examples - https://www.youtube.com/watch?v=qp1aOA4OD6E
4. KNIME Work Flow Construction
    a. Read Iris data set
    b. Train a classification model
    c. Test
    d. Score
5. KNIME Learning Resources - https://www.KNIME.com/resources

KNIME Model Deployment

1. Deploy PMML models as Rest API using OpenScoring
2. Run KNIME workflow in batch mode with a Listener node for input file
3. Run KNIME Server – Commercial component (USD 30K per year License fee)

KNIME Model Post-Deployment

1. Use Model Process Factory for monitoring deployed models and re-training when performance dips below a threshold

How to identify potential “learning from data” problem areas in your accounts

- Look for processes where decision making is happening by analyzing unstructured input data – Text, Image etc
- Absence of business rules relating inputs and required output but availability of historical data relating inputs and output.
- In the business processes, identify where SME’s engage in Exception processing. That is, transactions which cannot be processed by the business rules driven apps usually get routed to a queue where SMEs process these using their experience.
- Look for tasks where the number of input attributes impacting the output are large

