Master in Artificial Intelligence
Topics of Intelligent Systems
Practical Work Assignment
2025/26

Title: *Intelligent System for Industrial Equipment Failure Diagnosis*
Weight: 60% of final grade
Format: Group work (3 students)
Deadline: Dec 14, 2025

1. Motivation

Modern factories rely on intelligent systems to monitor equipment, anticipate failures, and recommend maintenance actions.
In this project, you will **combine probabilistic reasoning and symbolic knowledge representation** to build a hybrid decision-support prototype for diagnosing failures in a Computer-Numerical-Control (CNC) milling machine's spindle system.

This work will help you understand how:

- **Bayesian Networks** can model uncertainty and infer hidden causes,
- **Knowledge Graphs** can represent domain knowledge, actions, and explanations,
- Together, they support **actionable and explainable decisions**.

2. Learning Goals

By completing this assignment, you should be able to:

1. Apply **Bayes Decision Theory**, **MLE**, and **Expectation–Maximization (EM)** to real data.
2. Design and train a **Bayesian Network** for root-cause diagnosis.
3. Build a small **ontology/knowledge graph** for the machine domain.
4. Integrate BN inference with KG queries to recommend **maintenance actions**.
5. Deliver results in a reproducible, documented, and explainable way.

3. Use Case

You are tasked with diagnosing **Spindle Overheat** events in a CNC milling station.

- **Observed sensor data** (temperature, vibration, coolant flow, etc.) is available.
- **Latent causes** (BearingWear, FanFault, CloggedFilter, CoolingEfficiency) are not directly measured but influence the sensors.
- **Maintenance procedures** (ReplaceBearing, CleanFilter, RepairFan, FlushCoolant) are documented with effort, cost, and risk.

Your system should:

1. Use a **Bayesian Network** to infer overheat probability and likely causes.
2. Query a **Knowledge Graph** to identify possible maintenance procedures.
3. Recommend the best **action** (continue, slow down, or schedule maintenance) by minimizing expected cost.

4. Dataset and Resources

You are provided with:

- **Telemetry and labels** (telemetry.csv, labels.csv, maintenance.csv).
- **Knowledge graph master data** (components.csv, causes.csv, symptoms.csv, procedures.csv).
- **Ontology in RDF/OWL** (ontology.ttl).
- **Starter Python code** with BN, KG client, and decision logic.

You may extend or adapt the dataset/ontology as needed.

5. Deliverables

1. **Presentation (40%)**
    - Introduction & motivation
    - BN design, parameter learning (MLE, EM), inference results
    - Ontology/KG design and queries
    - Integration: how BN and KG work together
    - Results & evaluation (metrics, decision quality)
    - Reflection (strengths, limitations, improvements)
2. **Prototype Code (40%)**
    - Scripts to train BN and run inference
    - Ontology/KG implementation with SPARQL queries
    - Integration pipeline (BN → KG → decision)
    - Example run with explanation output (e.g.: *"Because Vibration=high and CoolantFlow=low, the system estimates Overheat=0.72. Likely cause is BearingWear. Recommended action: ReplaceBearing (6h, 600€)."*)
3. **Presentation & Demo (20%)**
    - 15-minute group presentation
    - Live demo of your system diagnosing a case
    - Questions and answers

6. Suggested tools

- Python with pgmpy, pandas, numpy
- RDF/OWL with rdflib or Protégé
- SPARQL queries with rdflib or SPARQLWrapper
- Streamlit or Jupyter for the demo interface (optional)

7. Summary of files content

File	Purpose
`data/telemetry.csv`	Machine sensor data — temperature, vibration, coolant, etc.
`data/labels.csv`	Whether the machine was overheated at that time (0 or 1).
`data/components.csv`	Components list for populating the KG.
`data/causes.csv`	Possible causes (BearingWear, FanFault...).
`data/symptoms.csv`	Observable symptoms (HighVibration, HighTemp...).
`data/relations.csv`	Links between them (e.g., Cause → Symptom).
`data/procedures.csv`	Maintenance actions and their effort/cost/risk.