arc4ml

Lightweight Architecture Documentation
for
Machine Learning Systems

12 Principles for Machine Learning Documentation

1. Vision & Business Process

2. Problem Description

3. Regulatory Framework & Risk

4. Assumptions, Constraints, & Special Cases

5. Adjacent Systems

6. Interfaces

7. System View

8. Model View

9. Data View

10. Experiments

11. Quality Aspects

12. Architectural Decisions

1
Vision & Business Process

Explain the business process and the expected business value.

Questions to Answer

- The 5 Why's?
- What is the expected business value?
- How is the business process structured?
- Which stakeholders are engaged?
- How can we measure the success (i.e., business metric?

Conceivable Sub-Chapters

- 1.1.Motivation
- 1.2. Description of Business Process
- 1.3. Success Criteria and Business Metrics
- 1.4. Stakeholder Map

2 Problem Description

Document how the business problem is framed as a machine learning formulation.

Questions to Answer

- How can we frame the business problem as a machine learning problem?
- Which technical metrics are relevant?
- Does the technical metric correlate with the business metric?

Conceivable Sub-Chapters

- 2.1. Problem Formulation
- 2.2. Technical Metrics

3 Regulatory Framework & Risk

Describe the regulatory forces acting upon the development process and potential risk originated from the ML system (due to wrong predictions).

Questions to Answer

- Which federal laws must be adhered to?
- Which organizational policies must be considered?
- Which risk grade must be assumed? (e.g., for EU AI Act)

Conceivable Sub-Chapters

- 3.1. Applicable Federal Laws
- 3.2. Organizational Regulations
- 3.3. Risk Assessment
- 3.4. Impact of Risk on Organization, Society, Humans

4
Assumptions, Constraints, & Special Cases

List and explain the assumptions and constraints that shape the development process.

Questions to Answer

- What assumptions were made?
- Which constraints apply?
- Which special cases must be considered?

Conceivable Sub-Chapters

- 4.1. Assumptions
- 4.2. Constraints
- 4.3. Special Cases

5
Adjacent Systems

Describe the adjacent software systems that must interact with the ML system.

Questions to Answer

- Which adjacent (i.e., neighbouring) systems are relevant?
- How are the adjacent systems related to the ML system at hand?
- Which business capabilities do the respective adjacent systems represent?

Conceivable Sub-Chapters

- 5.1. Overview of Adjacent Systems and respective Business Capabilities
- 5.2. Technical Context Diagram
- 5.3. Business Context Diagram

6 Interfaces

Describe how the ML system exchanges information and data with other systems.

Question to Answer

- Which interfaces exist?
- To which system and team do these interfaces belong?
- How do other systems and users engage with our system's predictions?
- How do we fetch data from and communicate with other systems?

Conceivable Sub-Chapters

- 6.1. Published Interfaces
  - 6.1.1. REST API
  - 6.1.2. Kafka Topic
  - 6.1.3. Cloud Storage
- 6.2. Consumed Interfaces
  - 6.2.1. REST API
  - 6.2.2. Kafka Topic
  - 6.2.3. Cloud Storage
  - 6.2.4. Databases

7
System View

Provide an overview of the ML system and its crucial components.

Question to Answer

- What is the tech stack?
- Which building blocks exist?
- How are the building blocks and the interfaces related?
- How does the system work (e.g., runtime)?
- Where do the respective components of the system reside (e.g., infrastructure, network)?

Conceivable Sub-Chapters

- 7.1. General Tech-Stack
- 7.2. Overview of Building Blocks
- 7.3. Overview of Runtime
- 7.4. Overview of Infrastructure
- 7.5. Overview of Deployment Pipeline

8
Model View

Describe the ML model in detail and the entire process for its lifecycle and all supporting building blocks.

Questions to Answer

- Which model (architecture) is used and why?
- When are model predictions made?
- Which features and hyperparameters are used?
- How is a ML model trained?
- When is a model (re-)trained?
- How are models deployed?
- How are models monitored?
- Where and how are versioned model artefacts stored?
- How to debug a deployed model?

Conceivable Sub-Chapters

- 8.1. Model Context
- 8.2. Model Runtime
- 8.3. Model Versioning
- 8.4. Model Training
- 8.4. Model Deployment
- 8.5. Model Monitoring
- 8.6. Model Configuration
- 8.7. Model Debugging

9
Data View

Provide a detailed overview of the data and features crucial for the ML model.

Questions to Answer

- Which and how much data is stored?
- How are data loaded and transformed?
- When and how often is data loaded?
- Is personal identifiable data used?
- How is data anonymized, if at all?
- How are the features calculated?
- Where are the features stored?
- Are the data and features versioned?

Conceivable Sub-Chapters

- 9.1. Data Description
- 9.2. Data Pipeline
- 9.3. Feature Engineering
- 9.4. Feature Store
- 9.5. Data Quality Tests
- 9.6. Data Storage and Retention
- 9.7. Data/Feature Versioning

10
Experiments

Explain how hypothesis are tested and provide an overview of past (computational) experiments.

Questions to Answer

- How to conduct computational experiments?
- Which hardware and software configurations must be used?
- Which hyperparameters/ models/features have been tested so far and what was the outcome?

Conceivable Sub-Chapters

- 10.1. Experimentation Setup
- 10.2. Experimentation Log

11
Quality Aspects

Outline the quality attributes of the ML system.

Questions to Answer

- Are the predictive results explainable?
- How to reproduce predictions?
- How susceptible is the ML system to small changes in the input?
- Are minorities/races/age groups disadvantaged by the system?
- How is the system secured?
- Are there technical debts?

Conceivable Sub-Chapters

- 10.1. Explainability
- 10.2. Reproducability
- 10.3. Robustness
- 10.4. Fairness
- 10.5. Scalability
- 10.6. Security
- 10.7. Logging
- 10.8. Technical Debt

12
Architectural Decisions

List all relevant architectural decisions made in the past.

Questions to Answer

- What architectural decisions have been made so far?
- And why?

Conceivable Sub-Chapters

- None

arc4ml

info: muhammed.demircan@arc4ml.com