Software Engineering for Artificial Intelligence

Requirements and Risks (Quality Assurance)

TECHNISCHE UNIVERSITÄT DARMSTADT

Outline

TECHNISCHE UNIVERSITÄT DARMSTADT

- Why Requirement Engineering for AI?
- Requirements Risk and Approaches

  Categories:

  - Technology
  - Regulations and Compliances
  - Human Resources
  - Model
  - Culture

Why Requirement Engineering for AI?

TECHNISCHE UNIVERSITÄT DARMSTADT

The Goal of AI Systems is to learn on Data and attempt to categorize or even predict the unknown, based on the acquired “knowledge”.

It is unavoidable that in many instances, it will fail at this task, no matter how well prepared the system is.

The question of approval of an AI based system should not be “does it make no mistakes?” but rather, “which kind of failures are we willing to accept and what other traits do we want to sacrifice to reduces the failure rate?”

Category: Technology - Requirements

TECHNISCHE UNIVERSITÄT DARMSTADT

AI heavily relies on Big Data by tons of Users to be (pre-)processed, covering their behavior, preferences and a lot of personal information, which are used in complex algorithms to make decisions

- User Data needs to be protected against adversary attacks/influence
- Prioritize comprehensibility of used algorithms to make detection of manipulation throughout the decision-making progress possible
- Ensure necessary infrastructure to process Big Data, the best system can not do its job when the underlying hardware can not handle the sheer mass of Input processing required in a reasonable time frame

Category: Technology - Risks

TECHNISCHE UNIVERSITÄT DARMSTADT

- User Data may be leaked or manipulated
- Manipulation attempts on results may go unnoticed when the complexity of the algorithm producing it is too high
- Deployment/Training of the AI System may be impossible due too lack of necessary hardware

Category: Technology – Approaches

TECHNISCHE UNIVERSITÄT DARMSTADT

- Avoid untrusted or unsupported open source software to make the search for weak points more difficult for attackers
- Be willing to sacrifice performance / accuracy to improve the understandability of the decision making process to notice and track changes better
- Invest in hardware infrastructure to handle big data processing to ensure the development and maintainability of the system

Category: Regulations and Compliances – Requirements

TECHNISCHE UNIVERSITÄT DARMSTADT

The Internet is a place where the complexity of the law and the consequence of breaking it are often difficult to comprehend. Even more when handling Big Data.

- Comply with data protection legislation and be aware of future changes to it
- Follow company data policies, make changes to it when necessary

Category: Regulations and Compliances – Risks

TECHNISCHE UNIVERSITÄT DARMSTADT

- Breaking the law…
- Making yourself legally vulnerable by violating the own terms of service
- System works on “obvious” illegal conditions
  - The gathering of specific data may be illegal, even if never used
- Required to start over the development/training of the system midway since it turns out one or more input parameter you are using is legally risky

Category: Regulations and Compliances – Approaches

TECHNISCHE UNIVERSITÄT DARMSTADT

Approaches:

- Stay in close contact with any legal advisors both for company intern matters and (inter)national laws
- Before the development phase, gather as much concrete information about the planned system as possible, even for training data.

Category: Human Resources – Requirements & Risks

TECHNISCHE UNIVERSITÄT DARMSTADT

Requirements:

- Defined roles & responsibilities throughout lifecycle of an AI
- AI Know-how employees needed even after deployment for adaptation to new obstacles

Risks:

- After deployment no one knows/wants to be responsible for changes of the system
- Third party AI skills during development no longer available
- Overdependence on too few AI savvy resources may lead to an unhealthy work distribution

Category: Human Resources – Approaches

TECHNISCHE UNIVERSITÄT DARMSTADT

Approaches:

- Define roles and their responsibilities which are required for development, maintenance and adaptation of the System
- Have AI savvy employees which are familiar with the system to be prepared for adaptations, be it fixing of existing features or coming up with new ones

Category: Model

TECHNISCHE UNIVERSITÄT DARMSTADT

Risk: Inaccuracy of results

- Incorrect type of algorithm(s) applied to a problem
- Poor data quality
- Suboptimal choice of algorithm parameters
- Inappropriate feedback going undetected (in those AI solutions allowing for continuous feedback and learning)

Category: Model

TECHNISCHE UNIVERSITÄT DARMSTADT

Risk: Client dissatisfaction

- Increased probability that business users may lack adequate understanding of complex AI model limitations and incorrectly interpret AI outputs

  ⇒ Risk of products being developed which do not meet customer needs

Category: Model

TECHNISCHE UNIVERSITÄT DARMSTADT

Requirement: Performance evaluation metrics

- Quality of the resulting predictions
- Good understanding of performance measures
  - Accuracy
  - Precision
  - Recall
  - Lift: factor by which the prediction of a model is more accurate than a random choice
- Appropriateness of measures depends on the problem domain.

Category: Model

TECHNISCHE UNIVERSITÄT DARMSTADT

Requirement: Explainability

- Explainable predictions of the model
- Elicit explainability requirements from a user's point view.
  - Model simplicity
    - Constrain the models to derive explanations.
    - Minimize the number of features.

Category: Model

TECHNISCHE UNIVERSITÄT DARMSTADT

Data requirements:

- Discuss conditions for data preparation, definitions of outliers, and derived data.
- Specify requirements regarding the collection of data, the data formats, and the ranges of data.
- Requirements on data quantity
  - Diversity of data
    - Identify additional data sources as part of the stakeholder analysis.

.

Category: Model

TECHNISCHE UNIVERSITÄT DARMSTADT

Data requirements:

- Requirements on data quality
  - Dimensions of data quality
    - Completeness: sparsity of data within each characteristic
    - Consistency: format and representation of data that should be the same in the dataset
    - Correctness: degree to which you can rely on the data actually being true

Category: Model

TECHNISCHE UNIVERSITÄT DARMSTADT

Data requirements:

- Requirements on data quality
  - Data provenance
    - Critically question the data sources.
    - Avoid the use of less trustable public datasets.
- Specify conditions for data anomalies.
- Monitor and analyze operational data.
- Retrain ML models.

Category: Culture

TECHNISCHE UNIVERSITÄT DARMSTADT

Risk: Discrimination

- Cultural challenge for large scale AI adoption due to actual or perceived regulatory and ethical concerns
- Discrimination is more implicit in ML systems than rule-based algorithms.
- ML algorithms amplify discrimination in the data during the training process.

Category: Culture

TECHNISCHE UNIVERSITÄT DARMSTADT

Requirement: Fairness

- Prepare the training data so that it does not contain any "protected" characteristics…
- … or analyze the trained model to find important features in the training data.
  - Assess whether these features may point to unacceptable discrimination.

Category: Culture

TECHNISCHE UNIVERSITÄT DARMSTADT

Requirement: Fairness

- Measurement error
  - Label bias: The outcome of interest may be imperfectly observed.
    - Focus on outcomes less likely to exhibit such bias.
  - Feature bias: The predictive power of features can vary across groups.
    - Minimize sample bias when constructing risk scores.
- Simple, interpretable, and explainable statistical models

References

- Vogelsang, Andreas, and Borg, Markus. *Requirements Engineering for Machine Learning: Perspectives from Data Scientists*. In Proc. of the 6th International Workshop on Artificial Intelligence for Requirements Engineering (AIRE). 2019. https://arxiv.org/pdf/1908.04674.pdf
- Rahimi, Mona et al. *Toward Requirements Specification for Machine-Learned Components*. 2019 IEEE 27th International Requirements Engineering Conference Workshops (REW). 2019. https://ieeexplore.ieee.org/document/8933771
- Hulten, Geoff. *Building Intelligent Systems: A Guide to Machine Learning Engineering*. Apress. 2018.
- Corbett-Davies, Sam, and Goel, Sharad. *The Measure and Mismeasure of Fairness:A Critical Review of Fair Machine Learning*. 2018. https://arxiv.org/pdf/1808.00023.pdf
- Christian Kaestner. *Machine Learning is Requirements Engineering—On the Role of Bugs, Verification, and Validation in Machine Learning*. 2020. https://medium.com/analytics-vidhya/machine-learning-is-requirements-engineering-8957aee55ef4
- Deloitte. *AI and risk management: Innovating with confidence*. 2018. https://www2.deloitte.com/bd/en/pages/financial-services/articles/gx-ai-and-risk-management.html

Acknowledgements & License

- These slides are made available by the authors (Valeria Marchan, Mathieu Hillen) under CC BY 4.0.