Machine Learning Operations Canvas (v1.1)

Product name: | **Designed by:** | **Date:** | **Iteration:**

Problem

Background

Describe the context, including the problem and business need. Explain why this ML project is important

Value Proposition

Outline the key benefits and the value the ML solution will bring. Highlight its impact on the business or users.

Objectives

State the specific, measurable goals of the ML project. Detail the expected outcomes and success criteria.

Data

Data Collection

Identify the data sources and methods for gathering data. Include information on data frequency, volume and labelling process.

Data Verification and Governance

Explain the data management policies, focusing on quality, privacy, and compliance. Include mechanisms for data access controls, quality checks, and compliance monitoring.

Model

Modelling

Detail the algorithms and techniques used for building the ML model. Include information on feature engineering and selection.

Metrics and Evaluation

Specify the performance metrics and evaluation methods. Describe how the model's effectiveness will be assessed.

Model Governance

Outline the process for managing models versions including conditions from going from staging to production. Outline procedures for updating and retraining models.

Operations

Inference

Describe the deployment process for the model to make predictions. Include details on the infrastructure and environment used.

Decision

Explain how the model's predictions are integrated into decision-making. Detail any human oversight or automated decision systems.

Monitoring

Feedback

Describe the mechanisms for collecting feedback on model performance. Explain how this feedback is used to refine the model.

Lifetime

Outline the lifetime after model deployment. This includes monitoring for model drift, conditions for retraining, and conditions for decommissioning.

Risk

Fairness

Evaluate potential biases in the data and model that could lead to unfair outcomes. Include strategies for identifying, measuring, and mitigating bias across the system.

Explainability

Detail how the model's decisions can be interpreted and understood by stakeholders. Include methods to enhance transparency and communicate decision-making processes effectively.

Security

Identify risks related to data breaches, adversarial attacks, and system vulnerabilities. Include measures for safeguarding data and ensuring model robustness against malicious exploitation.

DTU

By Nicki Skafte Detlefsen nsde@dtu.dk
From DTU course 02476 Machine Learning Operations

Machine Learning Operations Canvas (v1.1)

Product name:	Designed by:	Date:	Iteration:

Explainer

Anecdotally, as many as 90% of machine learning models never make it to production. While the reasons vary—from misaligned goals to operational bottlenecks—the message is clear: building a model is just one part of the journey. To truly succeed, teams must navigate the complexities of deployment, scaling, and monitoring, all while addressing ethical and strategic considerations.

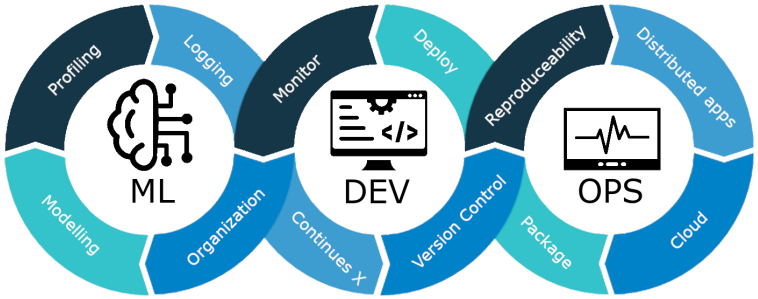
A machine learning model is only as impactful as the system that supports it — strategy, *scalability, and responsibility form its true foundation.*

The Machine Learning Operations (MLOps) Canvas is a structured framework designed to aid in the planning, execution, and management of machine learning projects. Drawing inspiration from the widely used Business Model Canvas [1], the MLOps Canvas aims to help teams consisting of both technical and non-technical members to collaborate on machine learning projects effectively. It is inspired by similar canvases [2-3], but with two key difference: more focus is placed on the operational aspects of machine and the risks associated with machine learning/AI projects [4].

The canvas is divided into six columns, representing critical stages of the machine learning lifecycle: **Problem, Data, Model, Operations, Monitoring, and Risk**. Each column is further divided into boxes that guide the user through the process. On the following pages, questions can be found for each box that are designed to spark discussions and help fill out the boxes. Some of the questions are technical, while others are more strategic or ethical in nature. Answering these is not easy, but investing time in this process early on can help reduce technical debt [5] in the future and increase the chances of project success [6]. Finally, the last page contains references for each box if you need to educate yourself further on a specific topic.

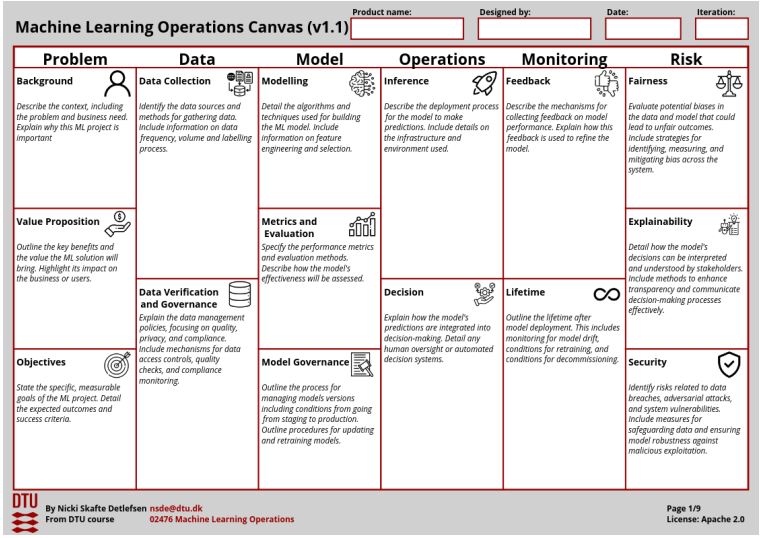
The canvas is designed to be used as a collaborative tool, allowing teams to work together to develop a shared understanding of the project and align on key decisions and objectives. It is intended to be a living document that evolves as the project progresses and new information becomes available. Therefore make sure to use the date and iteration fields to keep track of changes and updates.

[1] Alexander Osterwalder (2004). The Business Model Ontology: A Proposition In A Design Science Approach
[2] Goku Mohandas (2023). Machine Learning Canvas, https://madewithml.com/courses/mlops/product-design/
[3] Louis Dorard (2015). Machine Learning Canvas, https://www.machinelearningcanvas.com/
[4] European Commission (2024). EU AI ACT, https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai
[5] D. Sculley et. al. (2015). Hidden Technical Debt in Machine Learning Systems
[6] Chip Huyen (2022). Designing Machine Learning Systems: An Iterative Process for Production-Ready Applications



The canvas should be filled from left to right, starting with the Problem column and moving through each subsequent column in order. This reflects the underlying project lifecycle and order of operations.

In addition to the canvas, we recommend teams to add a project overview describing the required team members, roles, responsibilities and a projected timeline with milestones.


Machine Learning Operations Canvas (v1.1)

Product name: | Designed by: | Date: | Iteration:

Problem	Data	Model	Operations	Monitoring	Risk
Background *Describe the context, including the problem and business need. Explain why this ML project is important*	**Data Collection** *Identify the data sources and methods for gathering data. Include information on data frequency, volume and labelling process.*	**Modelling** *Detail the algorithms and techniques used for building the ML model. Include information on feature engineering and selection.*	**Inference** *Describe the deployment process for the model to make predictions. Include details on the infrastructure and environment used.*	**Feedback** *Describe the mechanisms for collecting feedback on model performance. Explain how this feedback is used to refine the model.*	**Fairness** *Evaluate potential biases in the data and model that could lead to unfair outcomes. Include strategies for identifying, measuring, and mitigating bias across the system.*
Value Proposition *Outline the key benefits and the value the ML solution will bring. Highlight its impact on the business or users.*		**Metrics and Evaluation** *Specify the performance metrics and evaluation methods. Describe how the model's effectiveness will be assessed.*			**Explainability** *Detail how the model's decisions can be interpreted and understood by stakeholders. Include methods to enhance transparency and communicate decision-making processes effectively.*
	Data Verification and Governance *Explain the data management policies, focusing on quality, privacy, and compliance. Include mechanisms for data access controls, quality checks, and compliance monitoring.*		**Decision** *Explain how the model's predictions are integrated into decision-making. Detail any human oversight or automated decision systems.*	**Lifetime** *Outline the lifetime after model deployment. This includes monitoring for model drift, conditions for retraining, and conditions for decommissioning.*	
Objectives *State the specific, measurable goals of the ML project. Detail the expected outcomes and success criteria.*		**Model Governance** *Outline the process for managing models versions including conditions from going from staging to production. Outline procedures for updating and retraining models.*			**Security** *Identify risks related to data breaches, adversarial attacks, and system vulnerabilities. Include measures for safeguarding data and ensuring model robustness against malicious exploitation.*

DTU By Nicki Skafte Detlefsen nsde@dtu.dk From DTU course 02476 Machine Learning Operations Page 1/9 License: Apache 2.0

Machine Learning Operations Canvas (v1.1)

Product name:	Designed by:	Date:	Iteration:

Problem

Clearly defining the problem ensures the project is addressing a meaningful challenge with a clear value proposition. By understanding the background, value, and objectives, stakeholders align on goals and avoid wasted effort on solutions that don't meet real-world needs. This foundation is crucial for keeping the project focused and outcome-driven.

Background

Describe the context, including the problem and business need. Explain why this ML project is important

⭐What is the context or environment in which this problem exists?
⭐Why is this problem worth solving?
⭐Who are the stakeholders affected by this problem?
⭐What is the current state of the problem?
⭐Are there any constraints or assumptions to consider?

Value Proposition

Outline the key benefits and the value the ML solution will bring. Highlight its impact on the business or users.

⭐What value will the solution provide to stakeholders?
⭐How does this solution improve upon existing alternatives?
⭐What specific needs or pain points does it address?
⭐Who benefits most from solving this problem?
⭐How will success be measured in terms of value delivered?

Objectives

State the specific, measurable goals of the ML project. Detail the expected outcomes and success criteria.

⭐What are the measurable goals of the machine learning system?
⭐How do these objectives align with the overall business or project goals?
⭐What are the short-term and long-term targets?
⭐How will you prioritize conflicting objectives, if any?
⭐Are there clear performance thresholds or benchmarks to achieve?

Machine Learning Operations Canvas (v1.1)

Product name:	Designed by:	Date:	Iteration:

Data

Data is the backbone of any machine learning project, and poor data quality or governance can leadto unreliable or biased models. Carefully planning data collection and verification ensures the system is built on accurate, representative, and trustworthy information, reducing risks and setting the project up for success.

Data Collection

Identify the data sources and methods for gathering data. Include information on data frequency, volume and labelling process.

- ⭐What are the primary sources of data, and how will the data be accessed?
- ⭐Is the data collected at a fixed frequency, real-time, or as a one-time batch?
- ⭐How much data is needed (volume), and how will it be labeled or annotated?
- ⭐Are there any licensing, ownership, or copyright issues related to the data?
- ⭐Does the data include diverse and representative samples to address the problem effectively?

Data Verification and Governance

Explain the data management policies, focusing on quality, privacy, and compliance. Include mechanisms for data access controls, quality checks, and compliance monitoring.

- ⭐What steps will you take to ensure data quality (e.g., handling missing, duplicate, or incorrect data)?
- ⭐How will you validate that the data accurately represents the real-world problem?
- ⭐Are there policies in place to govern the storage, access, and usage of the data?
- ⭐How will you handle sensitive data or ensure compliance with privacy laws (e.g., GDPR)?
- ⭐What methods will you use to detect and mitigate potential biases in the data?

Machine Learning Operations Canvas (v1.1)

Product name:	Designed by:	Date:	Iteration:

Model

The model is the engine of the machine learning system, and its design, evaluation, and governance directly affect the project's effectiveness and reliability. Planning this column ensures the model is not only high-performing but also compliant, robust, and aligned with the project's goals and constraints.

Modelling

Detail the algorithms and techniques used for building the ML model. Include information on feature engineering and selection.

- ⭐What algorithms or techniques will be used to build the model?
- ⭐How will you decide which features to include or exclude?
- ⭐What is the expected complexity of the model, and is it justified?
- ⭐What tools and frameworks will you use to develop the model?
- ⭐What trade-offs (e.g., accuracy vs. interpretability) will you need to consider?

Metrics and Evaluation

Specify the performance metrics and evaluation methods. Describe how the model's effectiveness will be assessed.

- ⭐What metrics will be used to evaluate model performance?
- ⭐How will you determine whether the model meets its objectives?
- ⭐What baseline or benchmarks will you compare the model against?
- ⭐How will you evaluate the model's performance across different data subsets?
- ⭐How will you validate the model's robustness to edge cases and noise?

Model Governance

Outline the process for managing models versions including conditions from going from staging to production. Outline procedures for updating and retraining models.

- ⭐What processes are in place to document the model's development lifecycle?
- ⭐How will you ensure accountability for model predictions?
- ⭐What safeguards are in place for monitoring ethical compliance?
- ⭐How will you track changes and versions of the model over time?
- ⭐Are there clear roles and responsibilities defined for maintaining the model?

Machine Learning Operations Canvas (v1.1)

Product name:	Designed by:	Date:	Iteration:

Operations

Planning how the model will operate in production ensures it delivers value effectively and reliably. By addressing inference and decision-making workflows, this column helps bridge the gap between technical implementation and real-world application, ensuring the project's outputs are actionable and integrated seamlessly.

Inference

Describe the deployment process for the model to make predictions. Include details on the infrastructure and environment used.

- ⭐How will the model's predictions be generated during production?
- ⭐What infrastructure is needed to support inference at scale?
- ⭐How will latency and throughput requirements be met?
- ⭐What methods will be used to handle errors or failed predictions?
- ⭐What format will the predictions be delivered in for downstream use?

Decision

Explain how the model's predictions are integrated into decision-making. Detail any human oversight or automated decision systems.

- ⭐How will predictions be integrated into decision-making workflows?
- ⭐Who or what will act on the model's predictions?
- ⭐Are there automated or manual checks in place for critical decisions?
- ⭐What are the expected impacts of these decisions on stakeholders?
- ⭐How will the model's outputs be communicated to end-users or stakeholders?

Machine Learning Operations Canvas (v1.1)

Product name:	Designed by:	Date:	Iteration:

Monitoring

Continuous monitoring is essential to maintain the model's performance and relevance over time. Addressing feedback mechanisms and lifecycle management in the planning stage ensures the project remains adaptable to change, preventing costly failures or outdated models in production.

Feedback

Describe the mechanisms for collecting feedback on model performance. Explain how this feedback is used to refine the model.

- ⭐What mechanisms are in place to collect feedback on the model's predictions?
- ⭐How will feedback be used to improve the model over time?
- ⭐Who will be responsible for analyzing and acting on the feedback?
- ⭐What channels will you use to gather user or system feedback?
- ⭐How will you address negative feedback or performance issues identified?

Lifetime

Outline the lifetime after model deployment. This includes monitoring for model drift, conditions for retraining, and conditions for decommissioning.

- ⭐What is the expected lifespan of the model before retraining or replacement?
- ⭐How will you monitor for model drift or performance degradation?
- ⭐What triggers will indicate the need for model updates?
- ⭐What processes are in place for retiring outdated models?
- ⭐How will you manage dependencies and compatibility over the model's lifetime?

Machine Learning Operations Canvas (v1.1)

Product name:	Designed by:	Date:	Iteration:

Risk

Machine learning systems can have far-reaching impacts, and addressing risks ensures they are ethical, fair, and secure. By focusing on fairness, explainability, and security, this column prepares projects to meet regulatory, societal, and organizational standards, reducing potential harm and building trust with stakeholders.

Fairness

Evaluate potential biases in the data and model that could lead to unfair outcomes. Include strategies for identifying, measuring, and mitigating bias across the system.

- ⭐What potential biases could arise in the data or model, and how will you identify them?
- ⭐How will you ensure equitable outcomes for all stakeholder groups?
- ⭐What metrics or methods will you use to measure fairness?
- ⭐What steps will you take to mitigate bias during development and deployment?
- ⭐Are there any groups or scenarios that may be disproportionately affected?

Explainability

Detail how the model's decisions can be interpreted and understood by stakeholders. Include methods to enhance transparency and communicate decision-making processes effectively.

- ⭐How will you ensure the model's decisions can be interpreted by stakeholders?
- ⭐What tools or techniques will you use to enhance explainability?
- ⭐How will you communicate the model's decision-making process to non-technical audiences?
- ⭐Are there trade-offs between explainability and performance to consider?
- ⭐What level of transparency is required for regulatory or ethical purposes?

Security

Identify risks related to data breaches, adversarial attacks, and system vulnerabilities. Include measures for safeguarding data and ensuring model robustness against malicious exploitation.

- ⭐What risks exist for data breaches or leaks, and how will you mitigate them?
- ⭐How will you safeguard against adversarial attacks on the model?
- ⭐What processes are in place to ensure data privacy during and after model usage?
- ⭐How will you monitor and address vulnerabilities in the deployed system?
- ⭐Are there compliance requirements related to security that must be addressed?

DTU
By Nicki Skafte Detlefsen nsde@dtu.dk
From DTU course 02476 Machine Learning Operations

References

Background

[1] Dominik Kreuzberger and Niklas Kühl and Sebastian Hirschl (2022). Machine Learning Operations (MLOps): Overview, Definition, and Architecture

[2] Shreya Shankar et al. (2022). Operationalizing Machine Learning: An Interview Study

[3] Meenu Mary John and Helena Holmström Olsson and Jan Bosch (Unknown Year). An empirical guide to MLOps adoption: Framework, maturity model and taxonomy

Value Proposition

[4] Payne, A. and Frow, P. and Eggert, A. (2017). The customer value proposition: evolution, development, and application in marketing

[5] Andreas Eggert et al. (2018). Conceptualizing and communicating value in business markets: From value in exchange to value in use

Objectives

[6] In Lee and Yongjae Shin (2020). Machine learning for enterprises: Applications, algorithm selection, and challenges

[7] Domingos, P. (2012). A few useful things to know about machine learning. Retrieved from https://doi.org/10.1145/2347736.2347755

[8] Andrew Ng (2024). Machine Learning Yearning. Retrieved from https://github.com/ajaymache/machine-learning-yearning

Data Collection

[9] Whang, S. E. et al. (2023). Data collection and quality challenges in deep learning: A data-centric ai perspective

[10] Yuji Roh and Geon Heo and Steven Euijong Whang (2018). A Survey on Data Collection for Machine Learning: A Big Data - AI Integration Perspective

[11] Alexandra L'Heureux et al. (2017). Machine Learning With Big Data: Challenges and Approaches

Data Verification and Governance

[12] V. Khatri and Carol V. Brown (2010). Designing data governance

[13] P. Malik (2013). Governing Big Data: Principles and practices

[14] Polyzotis, N. et al. (2019). Data validation for machine learning

Modelling

[15] Google Research (2024). Tuning Playbook. Retrieved from https://github.com/google-research/tuning_playbook

[16] {ML Contests} (2023). State of Competitive Machine Learning 2023. Retrieved from https://mlcontests.com/state-of-competitive-machine-learning-2023/

Metrics and Evaluation

[17] S. Raschka (2018). Model Evaluation, Model Selection, and Algorithm Selection in Machine Learning

[18] Ben Hutchinson et al. (2022). Evaluation Gaps in Machine Learning Practice

Model Governance

[19] A. Bedoya et al. (2022). A framework for the oversight and local deployment of safe and high-quality prediction models

[20] Jasmine Latendresse et al. (2024). An Exploratory Study on Machine Learning Model Management

[21] Sebastian Schelter et al. (2018). On Challenges in Machine Learning Model Management

Inference

[22] Nicolas Brousse (2024). Scaling Machine Learning Inference. Retrieved from https://nicolas.brousse.info/blog/scaling-machine-learning-inference/

[23] Wu, C. et al. (2019). Machine learning at facebook: Understanding inference at the edge

Decision

[24] Rahul Rai et al. (2021). Machine learning in manufacturing and industry 4.0 applications

[25] Jayatilake, S. M. D. A. C. and Ganegoda, G. U. (2021). Involvement of machine learning tools in healthcare decision making

Feedback

[26] Eric Breck et al. (2017). The ML test score: A rubric for ML production readiness and technical debt reduction

[27] Paleyes, A. and Urma, R. and Lawrence, N. D. (2022). Challenges in deploying machine learning: a survey of case studies

Lifetime

[28] Mallick, A. et al. (2022). Matchmaker: Data Drift Mitigation in Machine Learning for Large-Scale Systems

[29] D. Silver and Qiang Yang and Lianghao Li (2013). Lifelong Machine Learning Systems: Beyond Learning Algorithms

Fairness

[30] Simon Caton and C. Haas (2020). Fairness in Machine Learning: A Survey

[31] Ninareh Mehrabi et al. (2019). A Survey on Bias and Fairness in Machine Learning

[32] S. Corbett-Davies and Sharad Goel (2018). The Measure and Mismeasure of Fairness: A Critical Review of Fair Machine Learning

Explainability

[33] Linardatos, P. and Papastefanopoulos, V. and Kotsiantis, S. (2020). Explainable ai: A review of machine learning interpretability methods

[34] Carvalho, D. V. and Pereira, E. M. and Cardoso, J. S. (2019). Machine learning interpretability: A survey on methods and metrics

Security

[35] Xue, M. et al. (2020). Machine learning security: Threats, countermeasures, and evaluations

[36] Kathrin Grosse et al. (2022). Machine Learning Security in Industry: A Quantitative Survey

[37] Nicolas Papernot et al. (2016). Towards the Science of Security and Privacy in Machine Learning