Product Engineering for Machine Learning: A Grey Literature Review

Isaque Alves
Faculty of Gama (FGA)
University of Brasília (UnB)
Brasília, Brazil
isaquealvesdl@gmail.com

Leonardo Leite, Paulo Meirelles
Free and Open Source Software Competence Center
Institute of Mathematics and Statistics (IME)
University of São Paulo (USP)
São Paulo, Brazil
{leofl,paulormm}@ime.usp.br

Carla Rocha
Faculty of Gama (FGA)
University of Brasília (UnB)
Brasília, Brazil
caguiar@unb.br

Supplementary Material

Our objective of using Grey Literature is to provide a timely and broad perspective on the emerging topic of Machine Learning product lifecycle practices and methods from the practitioner's perspective. For analyzing the selected material content, we employed techniques from Grounded Theory (GT) [42].

Our study workflow has three phases: planning, data selection, and data analysis. We defined two simple query strings: (1) "ML product" and (2) "machine learning products". Some blogs also employ tags to categorize publications and facilitate the search. Therefore, in addition to the query strings, we incorporated the following tags: *product management*, *MLOps*, *artificial intelligence*, and *AI*.

The protocol for our first selection round, regarding quality criteria, was partly based on the guidelines suggested by Vahid *et al.* [43]. The selection criteria of this first round were:

- **Reputation:** we discarded works whose authors did not have experience in the area or other works published in the field. However, we considered works published by reputable organizations.
- **Methodology:** we discarded works without clear objectives and methodology.
- **Objectivity:** we discarded works with potential business interests; we considered only works with conclusions supported by data.

The protocol for our second selection round focused on the scope of the articles so that we could select only the essays pertinent to our objectives. Initially, we examined the titles, followed by the contents, and we favored the selection of articles that:

- focus on the development of ML products;
- discuss peculiarities and challenges for developing or managing ML products;
- report the use of practices applied to ML products; and
- discuss the ML product lifecycle, process, or workflow;

At the same time, we discarded posts:

- addressing some topic (e.g., ML in user experience) with ML products mentioned only for the background contextualization; and
- focusing on a single ML tool, without contextualizing it in ML product lifecycle.

We explored 10 data sources. Applying the planned procedures for data source selection, we chose two of them: Medium[1] and Toward Data Science[2] blogs, which have large communities of 532 thousand of readers and writers about our topic of interest. We discarded the following blogs because they are strictly focused on technical issues: Distill, BAIR Berkeley, Open AI, DeepMind Blog, and Colah's Blog. We rejected the following blogs in our GLR because they were related to private companies: Facebook AI's Blog, Google AI Blog, and Amazon AWS Machine Learning Blog. Finally, we discarded the blog Machine Learning at MIT because it discusses research results.

For the coding process [44], we highlighted excerpts linked to our research question, deriving from them practices and methods for ML product engineering. We constantly compared [44] the coding produced from different articles, and grouped common practices and methods, even when different posts used different terminology. In the following, we present some examples of extracted practices and methods, alongside the original excerpts associated with them:

- **Define the desired outcome** - "*define the objective function (outcome) and metrics...*" [1];
- **Review the literature** - "*The review of related literature (RRL) step of the ML workflow involves reading up on existing approaches, datasets, and others resources*" [22];
- **Data Requirement** - "*Figuring out what data are needed for a specific product or feature is the first and most important step in scoping data requirements.*" [23]).

From the GLR conducted in this study, we mapped 58 practices and methods related to the ML product development process. Table I presents the methods and practices emerged, and we grouped them into the following categories: *Problem definition and solution design*; *Product management*; *Data management*; *Model management*; *Software management*; *Delivery and runtime*.

[1] https://medium.com/
[2] https://towardsdatascience.com/

TABLE I
METHODS AND PRACTICES PRESENT ON ML PRODUCT LIFE CYCLE ORGANIZED BY CATEGORY.

Categories	Methods and Practices
Problem Definition and Solution Design	Business Continuous Validation [1]–[4]
	Verify how necessary is ML for the Product[1], [3], [5]–[14]
	Define the Role of ML on Product [4], [9]–[11], [14]–[16]
	Statement of Expectation and Intention [1], [3], [4], [9]–[12], [17]
	Build the Product Trust [1]
	Design Thinking [17]–[19]
	Lean Canvas [20]
	Prototyping [11], [18], [21]
	Define the Desired Outcome [1]–[3], [7], [11], [12], [22], [22], [23]
Product Management	Improvement Using User Feedback [9], [10], [13], [15], [23], [24]
	Establish what is the outcome and what the data can offer [6], [10]
	Review the Literature [3], [10], [22]
	Learn From Retrospective Meetings and Logs [6], [15]
	Balanced Scorecard [2]
	Risk Management [2], [13], [22]
	MLOps [25]–[30]
	CRISP-DM [31]–[33]
	Documentation [14], [34]
	Team Data Science Process (TDSP) [33]
	Multiple Interactions with users and stakeholders to colect feedback [3], [10]
	A/B testing or Split Testing [10], [14], [22], [24], [25]
	Evaluate Results, Define Metrics and Baselines [1], [3], [6], [7], [11], [12], [14], [15], [22]
	Agile Practices [4], [20], [24]
	DevOps [25], [27], [28], [30], [35]
	Define the Data Strategy [1], [2], [6], [11], [23], [36]
	Feedback Loops [4], [8], [11]–[13], [15], [37]
Data Management	Data Requirements [11], [23], [26], [36]
	Ensure the Reliability and Availability of Data [1], [6], [11], [12], [14], [23], [26]
	Define the Data Pipeline [1], [26], [34], [38], [39]
	Data Collection and Evolution [2], [4], [8], [11]–[13], [15], [16], [21]–[23], [26], [27], [34], [38]–[41]
	Data Cleaning [6]–[8], [11], [12], [14], [15], [21], [26]
	Data Labeling [8], [12], [22], [23], [26]
	Data Integrations [26]
	Data Versioning [4], [14], [15], [26], [34], [41]
	Data Transformation [26], [39]
	Data Reuse [26], [37]
Model Management	Research ML Libraries and Frameworks to be Used [6], [8], [11], [12], [17], [26], [38]
	Model Requirements [8], [12]
	Test Multiple Hypotheses [11], [14]
	Model Training [4], [11], [12], [14]–[16], [22], [27], [34], [39], [41]
	Measure Precision, Recall, and Accuracy [1], [8], [11], [17], [34]
	Model Evaluation [4], [12], [14]–[16], [22], [23], [27], [34], [37], [39], [40]
	Feature Engineering [7], [12], [15], [22], [27], [38], [41]
Software Management	Test Early and Frequently from end to end [1], [7], [11], [12]
	Code Reusability [6], [22], [26], [34]
	Modularizing Train Code [14]
	Model Versioning [14], [26], [28]
	Ensemble Learning [11], [14]
Delivery and Runtime	Model Deployment [22], [26], [39]
	Build Pipelines Specialized [6], [14], [15], [34], [41]
	Automation [15], [17], [22], [28], [37], [38]
	Continuous Testing [16], [27]
	Continuous Improvement [25], [30]
	Continuous Learning [19], [25], [27], [35], [37]
	Focus on Infrastructure [7], [34]
	CI/CD [29], [30]
	Continuous Model Monitoring [6], [13]–[15], [38]
	Continuous Success Measures [11], [27], [28]

REFERENCES

[1] B. Huang, "How to manage machine learning products." Blog post at Towards Data Science https://towardsdatascience.com/how-to-manage-machine-learning-products-part-1-386e7011258a, 2019.

[2] V. Patha, "A product manager's guide to machine learning: Balanced scorecard." Blog post at Towards Data Science https://towardsdatascience.com/a-product-managers-guide-to-machine-learning-balanced-scorecard-344cd37ab4a7, 2020.

[3] A. Mukherjee, "Product definition in the age of ai." Blog post at Towards Data Science https://towardsdatascience.com/product-definition-in-the-age-of-ai-619a417e3415, 2018.

[4] A. Pinhasi, "Machine learning products are built by teams, not unicorns." Blog post at Medium https://medium.com/@assaf.pinhasi/machine-learning-products-are-built-by-teams-not-unicorns-f3420c0e770b, 2019.

[5] B. Huang, "Three questions every ml product manager must answer." Blog post at Towards Data Science https:

//towardsdatascience.com/three-questions-every-ml-product-manager-must-answer-35d73127cd5d, 2020.
[6] C. I., "My framework for helping startups build and deploy data science." Blog post at Towards Data Science https://towardsdatascience.com/my-framework-for-helping-startups-build-and-deploy-data-science-43cf40bc1a1e, 2020.
[7] E. Dezhic, "Machine learning rules in a nutshell." Blog post at Towards Data Science https://towardsdatascience.com/machine-learning-rules-in-a-nutshell-dfc3e6839163, 2018.
[8] U. Barlaskar, "Machine learning for product managers." Blog post at Towards Data Science https://towardsdatascience.com/machine-learning-for-product-managers-part-i-problem-mapping-5436132c3a6e, 2018.
[9] G. Chandrasekhar, "7 elements of ai product strategy." Blog post at Towards Data Science https://towardsdatascience.com/defining-your-ai-product-strategy-7-areas-of-focus-2cf112c82c07, 2020.
[10] A. Mukherjee, "Lean startup and machine learning." Blog post at Towards Data Science https://towardsdatascience.com/lean-startup-and-machine-learning-39a96efecb2a, 2018.
[11] Y. Gavish, "The step-by-step pm guide to building machine learning based products." Blog post at Medium https://medium.com/@yaelg/product-manager-pm-step-by-step-tutorial-building-machine-learning-products-ffa7817aa8ab, 2017.
[12] J. Hurlock, "Building machine learning based products." Blog post at Medium https://medium.com/@jonhurlock/building-machine-learning-based-products-part-i-an-introduction-to-ml-products-406235e53be3, 2016.
[13] P. Pathak, "A product management framework for machine learning." Blog post at Medium https://productcoalition.com/my-framework-for-machine-learning-products-part-1-of-3-2354a0da63d2, 2018.
[14] S. Fabri, "Creating ml ai products — the smart products pipeline." Blog post at Medium https://medium.com/@simon.fabri_25096/the-smart-products-pipeline-404472842bfb, 2020.
[15] R. Agarwal, "Six important steps to build a machine learning system." Blog post at Towards Data Science https://towardsdatascience.com/6-important-steps-to-build-a-machine-learning-system-d75e3b83686, 2019.
[16] K. Schöfegger, "Lessons from building ml based products." Blog post at Medium https://medium.com/@kschoefegger/lessons-from-building-ml-based-products-4c7579724c91, 2019.
[17] Synthesized, "Shipping disruptive ml/ai products at scale (not papers). the engineer's take on scaling ml development processes." Blog post at Medium https://medium.com/@synthesized/shipping-disruptive-ml-ai-products-at-scale-not-papers-4008c6131f12, 2020.
[18] J. Rodrigues, "What's the role of designers while designing ai and ml products?." Blog post at Medium https://medium.com/@rodriguesjohnbaptist/whats-the-role-of-designers-while-designing-ai-and-ml-products-70e344285175, 2020.
[19] B. Gottipati, "Bringing design thinking to ai products." Blog post at Medium https://blog.usejournal.com/bringing-design-thinking-to-ai-products-f17386cf2ff2, 2019.
[20] D. Lianoudakis, "The ai product owner's tool kit." Blog post at Medium https://medium.com/@dlianou/the-ai-product-owners-tool-kit-aeaa2356f716, 2020.
[21] O. Schoppe, "The role of design in machine learning." Blog post at Medium https://medium.com/salesforce-ux/the-role-of-design-in-machine-learning-ae968ea90aac, 2020.
[22] L. Ampil, "Basics of data science product management: The ml workflow." Blog post at Towards Data Science https://towardsdatascience.com/basics-of-ai-product-management-orchestrating-the-ml-workflow-3601e0ff6baa, 2019.
[23] C. Mewald, "Data: A key requirement for your machine learning (ml) product." Blog post at Medium https://medium.com/thelaunchpad/data-a-key-requirement-for-your-machine-learning-ml-product-9195ace977d4, 2018.
[24] H. Harush, "Behind the buzzwords: how we build ml products at booking.com." Blog post at Medium https://medium.com/booking-product/behind-the-buzzwords-how-we-build-ml-products-at-booking-com-8140f8e47533, 2019.
[25] V. Karbhari, "Ml ops: Real world model deployments." Blog post at Medium https://medium.com/acing-ai/ml-ops-real-world-model-deployments-afab1fff9361, 2020.
[26] K. Fanous, "Software 2.0 vs software 1.0." Blog post at Medium https://medium.com/@karimfanous/software-2-0-vs-software-1-0-5df3f0d6d62a, 2020.
[27] A. Inc, "Machine learning pipeline: Architecture of ml platform in production." Blog post at Medium https://altexsoft.medium.com/machine-learning-pipeline-architecture-of-ml-platform-in-production-70e38571a89a, 2020.
[28] A. Soni, "Mlops: The epoch of productionizing ml models." Blog post at Medium https://medium.com/analytics-vidhya/mlops-the-epoch-of-productionizing-ml-models-4eec06d93623, 2020.
[29] A. Dhinakaran, "Ml infrastructure tools for production." Blog post at Towards Data Science https://towardsdatascience.com/ml-infrastructure-tools-for-production-1b1871eecafb, 2020.
[30] Y. Haviv, "How to productize ml faster with mlops automation." Blog post at Towards Data Science https://towardsdatascience.com/how-to-productize-ml-faster-with-mlops-automation-d3f25caf38c4, 2020.
[31] R. Sharma, "Hey airbnb, do you want more dollar bills?." Blog post at Towards Data Science https://towardsdatascience.com/crisp-dm-process-for-data-analysis-ed89db32cc88, 2019.
[32] R. S. Vuppuluri, "Yet another full stack data science project." Blog post at Towards Data Science https://towardsdatascience.com/yet-another-full-stack-data-science-project-a-crisp-dm-implementation-2943cae8da09, 2019.
[33] I. Rodrigues, "How to run a data science team: Tdsp and crisp methodologies." Blog post at Towards Data Science https://medium.com/skim-technologies/how-to-run-a-data-science-team-tdsp-and-crisp-methodologies-13bec343b406, 2019.
[34] P. Ataee, "Build an ml product — 4 mistakes to avoid." Blog post at Towards Data Science https://towardsdatascience.com/build-an-ml-product-4-mistakes-to-avoid-bce30d98bd24, 2020.
[35] M. Kumar, "A guide to deploying machine/deep learning model(s) in production." Blog post at Medium https://blog.usejournal.com/a-guide-to-deploying-machine-deep-learning-model-s-in-production-e497fd4b734a, 2018.
[36] C. Mewald, "Data: A key requirement for your machine learning (ml) product." Blog post at Medium https://medium.com/thelaunchpad/data-a-key-requirement-for-your-machine-learning-ml-product-9195ace977d4, 2018.
[37] C. Mewald, "3 common problems with your machine learning product and how to fix them." Blog post at Medium https://medium.com/thelaunchpad/how-to-protect-your-machine-learning-product-from-time-adversaries-and-itself-ff07727d6712, 2018.
[38] N. Lathia, "Machine learning for product managers." Blog post at Medium https://medium.com/@neal_lathia/machine-learning-for-product-managers-ba9cf8724e57, 2017.
[39] C. Mewald, "Your deep-learning-tools-for-enterprises startup will fail." Blog post at Medium https://medium.com/thelaunchpad/your-deep-learning-tools-for-enterprises-startup-will-fail-94fb70683834, 2019.
[40] R. Raghuwanshi, "Building ml products: Common pitfalls that product managers should avoid." Blog post at Medium https://medium.com/@rahulraghuwanshi_13455/building-ml-products-common-pitfalls-that-product-managers-should-avoid-14f393d1c250, 2019.
[41] J. Cogan, "The surprising truth about what it takes to build a machine learning product." Blog post at Medium https://medium.com/thelaunchpad/the-ml-surprise-f54706361a6c, 2019.
[42] K. Charmaz, "Grounded theory as an emergent method," *Handbook of emergent methods*, vol. 155, p. 172, 2008.
[43] V. Garousi, M. Felderer, and M. V. Mäntylä, "Guidelines for including grey literature and conducting multivocal literature reviews in software engineering," *Information and Software Technology*, vol. 106, pp. 101–121, 2019.
[44] B. Glaser and A. Strauss, *The discovery of grounded theory: strategies for qualitative research*. Aldine Transaction, 1999.