

Machine Learning Factory

From the Idea to the Continuous Maintenance of any
Machine Learning Application

Larysa Visengeriyeva

Machine Learning Factory

From the Idea to the Continuous Maintenance of any
Machine Learning Application

Larysa Visengeriyeva

This book is for sale at <http://leanpub.com/machinelearningfactory>

This version was published on 2020-09-21



This is a [Leanpub](#) book. Leanpub empowers authors and publishers with the Lean Publishing process. [Lean Publishing](#) is the act of publishing an in-progress ebook using lightweight tools and many iterations to get reader feedback, pivot until you have the right book and build traction once you do.

© 2020 Larysa Visengeriyeva

Contents

Introduction	1
Why is MLE different from Software Engineering	1
Business and Data Understanding	2
Defining the Scope of the Machine Learning Application	2
Requirements Engineering for Machine Learning	2
ML Project Management	2
Data Understanding and Collection	2
Labs	3
References	3
Data Preparation	4
Preparing Data for ML	4
Clean Data	4
Constructing Data	5
DataOps for ML	5
Labs	6
References	6
Modeling	7
Research on similar ML problems	7
Defining Quality Metrics of ML Model	7
Model Selection and Baseline Specification	8
Domain Knowledge Incorporation for Model Quality Improvement	8
Model Training	8
Model Compression	9
Ensemble Methods	9
Assuring Model Training Reproducibility	9
ML Training Pipelines	10
Labs	10
References	10
Evaluation	11
Performance Validation	11
Robustness Determination	11

CONTENTS

Increasing Explainability	11
Labs	11
References	11
Deployment	12
Inference Hardware Specification	12
Model Evaluation Under Production Condition	12
Versioning for ML Applications	12
CI/CD for ML	12
Deployment Strategy (Canary Deployment)	12
Model Serving Patterns	12
Fall-back Plan	13
Labs	13
References	13
Conclusion	14

Introduction

This is an introduction into the Machine Learning Engineering field.

Why is MLE different from Software Engineering

Data

Model

Code

Business and Data Understanding

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/machinelearningfactory>.

Defining the Scope of the Machine Learning Application

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/machinelearningfactory>.

Defining the Success Criteria for Business and ML Aspects

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/machinelearningfactory>.

Feasibility of ML Project

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/machinelearningfactory>.

Requirements Engineering for Machine Learning

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/machinelearningfactory>.

ML Project Management

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/machinelearningfactory>.

Data Understanding and Collection

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/machinelearningfactory>.

Data Estimation

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/machinelearningfactory>.

Data Versioning

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/machinelearningfactory>.

Data Quality Verification

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/machinelearningfactory>.

Labs

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/machinelearningfactory>.

References

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/machinelearningfactory>.

Data Preparation

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/machinelearningfactory>.

Preparing Data for ML

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/machinelearningfactory>.

Feature Selection

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/machinelearningfactory>.

Data Selection

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/machinelearningfactory>.

Dealing with Unbalanced Classes

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/machinelearningfactory>.

Feature Stores

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/machinelearningfactory>.

Clean Data

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/machinelearningfactory>.

Error Detection/Correction

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/machinelearningfactory>.

Data Wrangling

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/machinelearningfactory>.

Data Imputation

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/machinelearningfactory>.

Constructing Data

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/machinelearningfactory>.

Feature Engineering

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/machinelearningfactory>.

Data Augmentation

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/machinelearningfactory>.

Data Consistency with Data Standardization

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/machinelearningfactory>.

DataOps for ML

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/machinelearningfactory>.

Labs

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/machinelearningfactory>.

References

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/machinelearningfactory>.

Modeling

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/machinelearningfactory>.

Research on similar ML problems

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/machinelearningfactory>.

Defining Quality Metrics of ML Model

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/machinelearningfactory>.

Performance

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/machinelearningfactory>.

Robustness

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/machinelearningfactory>.

Scalability

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/machinelearningfactory>.

Explainability

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/machinelearningfactory>.

Model Complexity

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/machinelearningfactory>.

Resource Demand

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/machinelearningfactory>.

Model Selection and Baseline Specification

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/machinelearningfactory>.

Domain Knowledge Incorporation for Model Quality Improvement

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/machinelearningfactory>.

Model Training

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/machinelearningfactory>.

Objective

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/machinelearningfactory>.

Optimizer

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/machinelearningfactory>.

Regularization

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/machinelearningfactory>.

Cross-validation

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/machinelearningfactory>.

Model Compression

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/machinelearningfactory>.

Ensemble Methods

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/machinelearningfactory>.

Assuring Model Training Reproducibility

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/machinelearningfactory>.

Method Reproducibility

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/machinelearningfactory>.

Result Reproducibility

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/machinelearningfactory>.

Experimental Documentation

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/machinelearningfactory>.

ML Training Pipelines

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/machinelearningfactory>.

Labs

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/machinelearningfactory>.

References

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/machinelearningfactory>.

Evaluation

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/machinelearningfactory>.

Performance Validation

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/machinelearningfactory>.

Robustness Determination

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/machinelearningfactory>.

Increasing Explainability

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/machinelearningfactory>.

Features Observation on Model Prediction

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/machinelearningfactory>.

Labs

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/machinelearningfactory>.

References

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/machinelearningfactory>.

Deployment

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/machinelearningfactory>.

Inference Hardware Specification

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/machinelearningfactory>.

Model Evaluation Under Production Condition

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/machinelearningfactory>.

Versioning for ML Applications

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/machinelearningfactory>.

CI/CD for ML

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/machinelearningfactory>.

Deployment Strategy (Canary Deployment)

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/machinelearningfactory>.

Model Serving Patterns

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/machinelearningfactory>.

Fall-back Plan

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/machinelearningfactory>.

Labs

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/machinelearningfactory>.

References

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/machinelearningfactory>.

Conclusion

This content is not available in the sample book. The book can be purchased on Leanpub at <http://leanpub.com/machinelearningfactory>.