visenger / awesome-mlops Public

Code Issues 3 Pull requests 1 Actions Projects Security Insights

master

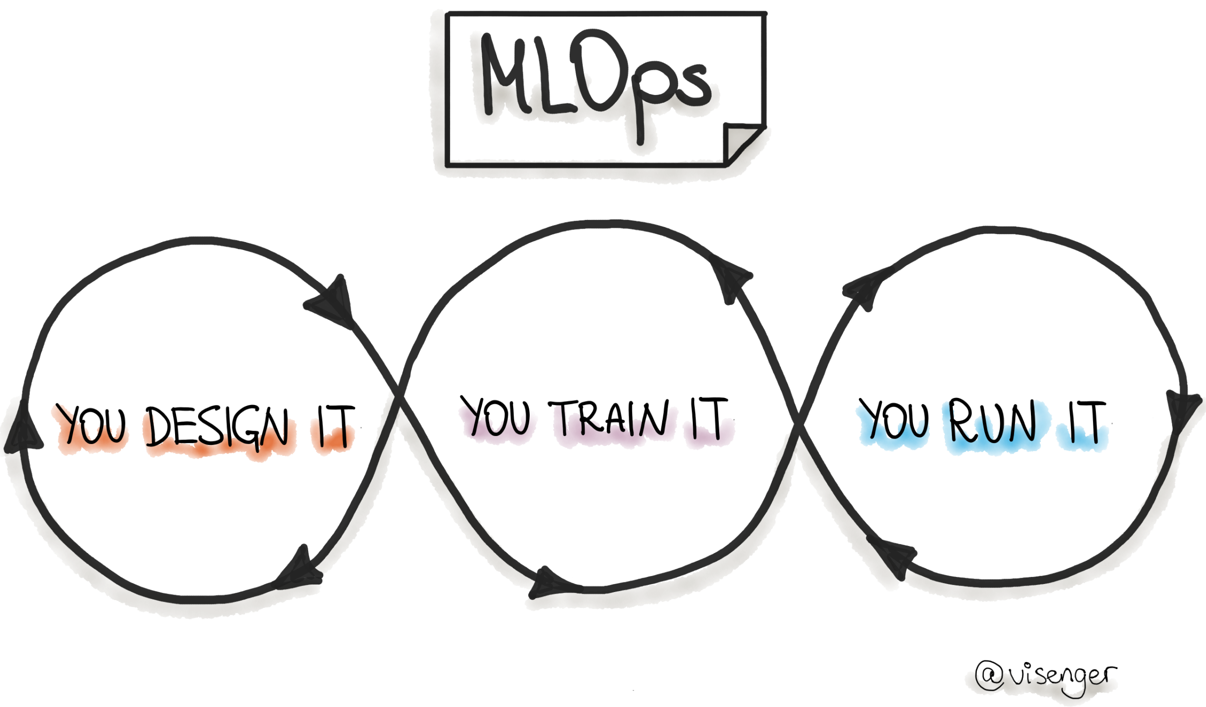
awesome-mlops / README.md

krishkatyal added author details

51 contributors +35

680 lines (606 sloc) 70.9 KB

Awesome MLOps awesome Made With Love



An awesome list of references for MLOps - Machine Learning Operations 👉 ml-ops.org

Follow @visenger 1.8k

Table of Content

MLOps Books	MLOps Articles
MLOps Workflow Management	MLOps: Feature Stores
MLOps: Data Engineering (DataOps)	MLOps: Model Deployment and Serving
MLOps: Testing, Monitoring and Maintenance	MLOps: Infrastructure
MLOps Papers	Talks About MLOps
Existing ML Systems	Machine Learning
Software Engineering	Product Management for ML/AI
The Economics of ML/AI	Model Governance, Ethics, Responsible AI
MLOps: People & Processes	Newsletters About MLOps, Machine Learning, Data Science and Co.

MLOps Core

▼ Click to expand!

1. Machine Learning Operations: You Design It, You Train It, You Run It!
2. MLOps SIG Specification
3. ML in Production
4. Awesome production machine learning: State of MLOps Tools and Frameworks
5. Udemy “Deployment of ML Models”
6. Full Stack Deep Learning
7. Engineering best practices for Machine Learning
8. 🚀 Putting ML in Production
9. Stanford MLSys Seminar Series
10. IBM ML Operationalization Starter Kit
11. Productize ML. A self-study guide for Developers and Product Managers building Machine Learning products.
12. MLOps (Machine Learning Operations) Fundamentals on GCP
13. ML full Stack preparation
14. MLOps Guide: Theory and Implementation
15. Practitioners guide to MLOps: A framework for continuous delivery and automation of machine learning.

MLOps Communities

▼ Click to expand!

1. MLOps.community
2. CDF Special Interest Group - MLOps
3. RsqrdAI - Robust and Responsible AI
4. DataTalks.Club
5. Synthetic Data Community
6. MLOps World Community

MLOps Courses

1. MLOps Zoomcamp (free)
2. Coursera's Machine Learning Engineering for Production (MLOps) Specialization

MLOps Books

▼ Click to expand!

1. "Machine Learning Engineering" by Andriy Burkov, 2020
2. "ML Ops: Operationalizing Data Science" by David Sweenor, Steven Hillion, Dan Rope, Dev Kannabiran, Thomas Hill, Michael O'Connell
3. "Building Machine Learning Powered Applications" by Emmanuel Ameisen
4. "Building Machine Learning Pipelines" by Hannes Hapke, Catherine Nelson, 2020, O'Reilly
5. "Managing Data Science" by Kirill Dubovikov
6. "Accelerated DevOps with AI, ML & RPA: Non-Programmer's Guide to AIOPS & MLOPS" by Stephen Fleming
7. "Evaluating Machine Learning Models" by Alice Zheng
8. Agile AI. 2020. By Carlo Appugliese, Paco Nathan, William S. Roberts. O'Reilly Media, Inc.
9. "Machine Learning Logistics". 2017. By T. Dunning et al. O'Reilly Media Inc.
10. "Machine Learning Design Patterns" by Valliappa Lakshmanan, Sara Robinson, Michael Munn. O'Reilly 2020
11. "Serving Machine Learning Models: A Guide to Architecture, Stream Processing Engines, and Frameworks" by Boris Lublinsky, O'Reilly Media, Inc. 2017
12. "Kubeflow for Machine Learning" by Holden Karau, Trevor Grant, Ilan Filonenko, Richard Liu, Boris Lublinsky
13. "Clean Machine Learning Code" by Moussa Taifi. Leanpub. 2020
14. E-Book "Practical MLOps. How to Get Ready for Production Models"
15. "Introducing MLOps" by Mark Treveil, et al. O'Reilly Media, Inc. 2020
16. "Machine Learning for Data Streams with Practical Examples in MOA", Bifet, Albert and Gavald`a, Ricard and Holmes, Geoff and Pfahringer, Bernhard, MIT Press, 2018
17. "Machine Learning Product Manual" by Laszlo Sragner, Chris Kelly

18. "Data Science Bootstrap Notes" by Eric J. Ma
19. "Data Teams" by Jesse Anderson, 2020
20. "Data Science on AWS" by Chris Fregly, Antje Barth, 2021
21. “Engineering MLOps” by Emmanuel Raj, 2021
22. Machine Learning Engineering in Action
23. Practical MLOps
24. "Effective Data Science Infrastructure" by Ville Tuulos, 2021
25. AI and Machine Learning for On-Device Development, 2021, By Laurence Moroney. O'Reilly
26. Designing Machine Learning Systems ,2022 by Chip Huyen , O'Reilly
27. Reliable Machine Learning. 2022. By Cathy Chen, Niall Richard Murphy, Kranti Parisa, D. Sculley, Todd Underwood. O'Reilly

MLOps Articles

▼ Click to expand!

1. Continuous Delivery for Machine Learning (by Thoughtworks)
2. What is MLOps? NVIDIA Blog
3. MLSpec: A project to standardize the intercomponent schemas for a multi-stage ML Pipeline.
4. The 2021 State of Enterprise Machine Learning | State of Enterprise ML 2020: PDF and Interactive
5. Organizing machine learning projects: project management guidelines.
6. Rules for ML Project (Best practices)
7. ML Pipeline Template
8. Data Science Project Structure
9. Reproducible ML
10. ML project template facilitating both research and production phases.
11. Machine learning requires a fundamentally different deployment approach. As organizations embrace machine learning, the need for new deployment tools and strategies grows.
12. Introducting Flyte: A Cloud Native Machine Learning and Data Processing Platform
13. Why is DevOps for Machine Learning so Different?
14. Lessons learned turning machine learning models into real products and services – O’Reilly
15. MLOps: Model management, deployment and monitoring with Azure Machine Learning
16. Guide to File Formats for Machine Learning: Columnar, Training, Inferencing, and the Feature Store
17. Architecting a Machine Learning Pipeline How to build scalable Machine Learning systems

18. Why Machine Learning Models Degrade In Production
19. Concept Drift and Model Decay in Machine Learning
20. Machine Learning in Production: Why You Should Care About Data and Concept Drift
21. Bringing ML to Production
22. A Tour of End-to-End Machine Learning Platforms
23. MLOps: Continuous delivery and automation pipelines in machine learning
24. AI meets operations
25. What would machine learning look like if you mixed in DevOps? Wonder no more, we lift the lid on MLOps
26. Forbes: The Emergence Of ML Ops
27. Cognilytica Report "ML Model Management and Operations 2020 (MLOps)"
28. Introducing Cloud AI Platform Pipelines
29. A Guide to Production Level Deep Learning
30. The 5 Components Towards Building Production-Ready Machine Learning Systems
31. Deep Learning in Production (references about deploying deep learning-based models in production)
32. Machine Learning Experiment Tracking
33. The Team Data Science Process (TDSP)
34. MLOps Solutions (Azure based)
35. Monitoring ML pipelines
36. Deployment & Explainability of Machine Learning COVID-19 Solutions at Scale with Seldon Core and Alibi
37. Demystifying AI Infrastructure
38. Organizing machine learning projects: project management guidelines.
39. The Checklist for Machine Learning Projects (from Aurélien Géron,"Hands-On Machine Learning with Scikit-Learn and TensorFlow")
40. Data Project Checklist by Jeremy Howard
41. MLOps: not as Boring as it Sounds
42. 10 Steps to Making Machine Learning Operational. Cloudera White Paper
43. MLOps is Not Enough. The Need for an End-to-End Data Science Lifecycle Process.
44. Data Science Lifecycle Repository Template
45. Template: code and pipeline definition for a machine learning project demonstrating how to automate an end to end ML/AI workflow.
46. Nitpicking Machine Learning Technical Debt
47. The Best Tools, Libraries, Frameworks and Methodologies that Machine Learning Teams Actually Use – Things We Learned from 41 ML Startups
48. Software Engineering for AI/ML - An Annotated Bibliography
49. Intelligent System. Machine Learning in Practice
50. CMU 17-445/645: Software Engineering for AI-Enabled Systems (SE4AI)
51. Machine Learning is Requirements Engineering

52. Machine Learning Reproducibility Checklist
53. Machine Learning Ops. A collection of resources on how to facilitate Machine Learning Ops with GitHub.
54. Task Cheatsheet for Almost Every Machine Learning Project A checklist of tasks for building End-to-End ML projects
55. Web services vs. streaming for real-time machine learning endpoints
56. How PyTorch Lightning became the first ML framework to run continuous integration on TPUs
57. The ultimate guide to building maintainable Machine Learning pipelines using DVC
58. Continuous Machine Learning (CML) is CI/CD for Machine Learning Projects (DVC)
59. What I learned from looking at 200 machine learning tools | Update: MLOps Tooling Landscape v2 (+84 new tools) - Dec '20
60. Big Data & AI Landscape
61. Deploying Machine Learning Models as Data, not Code — A better match?
62. "Thou shalt always scale" — 10 commandments of MLOps
63. Three Risks in Building Machine Learning Systems
64. Blog about ML in production (by maiot.io)
65. Back to the Machine Learning fundamentals: How to write code for Model deployment. Part 1, Part 2, Part 3
66. MLOps: Machine Learning as an Engineering Discipline
67. ML Engineering on Google Cloud Platform (hands-on labs and code samples)
68. Deep Reinforcement Learning in Production. The use of Reinforcement Learning to Personalize User Experience at Zynga
69. What is Data Observability?
70. A Practical Guide to Maintaining Machine Learning in Production
71. Continuous Machine Learning. Part 1, Part 2. Part 3 is coming soon.
72. The Agile approach in data science explained by an ML expert
73. Here is what you need to look for in a model server to build ML-powered services
74. The problem with AI developer tools for enterprises (and what IKEA has to do with it)
75. Streaming Machine Learning with Tiered Storage
76. Best practices for performance and cost optimization for machine learning (Google Cloud)
77. Lean Data and Machine Learning Operations
78. A Brief Guide to Running ML Systems in Production Best Practices for Site Reliability Engineers
79. AI engineering practices in the wild - SIG | Getting software right for a healthier digital world
80. SE-ML | The 2020 State of Engineering Practices for Machine Learning
81. Awesome Software Engineering for Machine Learning (GitHub repository)
82. Sampling isn't enough, profile your ML data instead

83. Reproducibility in ML: why it matters and how to achieve it
84. 12 Factors of reproducible Machine Learning in production
85. MLOps: More Than Automation
86. Lean Data Science
87. Engineering Skills for Data Scientists
88. DAGsHub Blog. Read about data science and machine learning workflows, MLOps, and open source data science
89. Data Science Project Flow for Startups
90. Data Science Engineering at Shopify
91. Building state-of-the-art machine learning technology with efficient execution for the crypto economy
92. Completing the Machine Learning Loop
93. Deploying Machine Learning Models: A Checklist
94. Global MLOps and ML tools landscape (by MLReef)
95. Why all Data Science teams need to get serious about MLOps
96. MLOps Values (by Bart Grasza)
97. Machine Learning Systems Design (by Chip Huyen)
98. Designing an ML system (Stanford | CS 329 | Chip Huyen)
99. How COVID-19 Has Infected AI Models (about the data drift or model drift concept)
100. Microkernel Architecture for Machine Learning Library. An Example of Microkernel Architecture with Python Metaclass
101. Machine Learning in production: the Booking.com approach
102. What I Learned From Attending TWIMLcon 2021 (by James Le)
103. Designing ML Orchestration Systems for Startups. A case study in building a lightweight production-grade ML orchestration system
104. Towards MLOps: Technical capabilities of a Machine Learning platform | Prosus AI Tech Blog
105. Get started with MLOps A comprehensive MLOps tutorial with open source tools
106. From DevOps to MLOPS: Integrate Machine Learning Models using Jenkins and Docker
107. Example code for a basic ML Platform based on Pulumi, FastAPI, DVC, MLFlow and more
108. Software Engineering for Machine Learning: Characterizing and Detecting Mismatch in Machine-Learning Systems
109. TWIML Solutions Guide
110. How Well Do You Leverage Machine Learning at Scale? Six Questions to Ask
111. Getting started with MLOps: Selecting the right capabilities for your use case
112. The Latest Work from the SEI: Artificial Intelligence, DevSecOps, and Security Incident Response
113. MLOps: The Ultimate Guide. A handbook on MLOps and how to think about it
114. Enterprise Readiness of Cloud MLOps
115. Should I Train a Model for Each Customer or Use One Model for All of My Customers?

116. MLOps-Basics (GitHub repo) by raviraja
117. Another tool won't fix your MLOps problems
118. Best MLOps Tools: What to Look for and How to Evaluate Them (by NimbleBox.ai)
119. MLOps vs. DevOps: A Detailed Comparison (by NimbleBox.ai)
120. A Guide To Setting Up Your MLOps Team (by NimbleBox.ai)

MLOps: Workflow Management

1. Open-source Workflow Management Tools: A Survey by Ploomber
2. How to Compare ML Experiment Tracking Tools to Fit Your Data Science Workflow (by dagshub)
3. 15 Best Tools for Tracking Machine Learning Experiments

MLOps: Feature Stores

▼ Click to expand!

1. Feature Stores for Machine Learning Medium Blog
2. MLOps with a Feature Store
3. Feature Stores for ML
4. Hopsworks: Data-Intensive AI with a Feature Store
5. Feast: An open-source Feature Store for Machine Learning
6. What is a Feature Store?
7. ML Feature Stores: A Casual Tour
8. Comprehensive List of Feature Store Architectures for Data Scientists and Big Data Professionals
9. ML Engineer Guide: Feature Store vs Data Warehouse (vendor blog)
10. Building a Gigascale ML Feature Store with Redis, Binary Serialization, String Hashing, and Compression (DoorDash blog)
11. Feature Stores: Variety of benefits for Enterprise AI.
12. Feature Store as a Foundation for Machine Learning
13. ML Feature Serving Infrastructure at Lyft
14. Feature Stores for Self-Service Machine Learning
15. The Architecture Used at LinkedIn to Improve Feature Management in Machine Learning Models.
16. Is There a Feature Store Over the Rainbow? How to select the right feature store for your use case

MLOps: Data Engineering (DataOps)

▼ Click to expand!

1. The state of data quality in 2020 – O'Reilly
2. Why We Need DevOps for ML Data
3. Data Preparation for Machine Learning (7-Day Mini-Course)
4. Best practices in data cleaning: A Complete Guide to Everything You Need to Do Before and After Collecting Your Data.
5. 17 Strategies for Dealing with Data, Big Data, and Even Bigger Data
6. DataOps Data Architecture
7. Data Orchestration — A Primer
8. 4 Data Trends to Watch in 2020
9. CSE 291D / 234: Data Systems for Machine Learning
10. A complete picture of the modern data engineering landscape
11. Continuous Integration for your data with GitHub Actions and Great Expectations. One step closer to CI/CD for your data pipelines
12. Emerging Architectures for Modern Data Infrastructure
13. Awesome Data Engineering. Learning path and resources to become a data engineer
14. Data Quality at Airbnb Part 1 | Part 2
15. DataHub: Popular metadata architectures explained
16. Financial Times Data Platform: From zero to hero. An in-depth walkthrough of the evolution of our Data Platform
17. Alki, or how we learned to stop worrying and love cold metadata (Dropbox)
18. A Beginner's Guide to Clean Data. Practical advice to spot and avoid data quality problems (by Benjamin Greve)
19. ML Lake: Building Salesforce's Data Platform for Machine Learning
20. Data Catalog 3.0: Modern Metadata for the Modern Data Stack
21. Metadata Management Systems
22. Essential resources for data engineers (a curated recommended read and watch list for scalable data processing)
23. Comprehensive and Comprehensible Data Catalogs: The What, Who, Where, When, Why, and How of Metadata Management (Paper)
24. What I Learned From Attending DataOps Unleashed 2021 (byJames Le)
25. Uber's Journey Toward Better Data Culture From First Principles
26. Cerberus - lightweight and extensible data validation library for Python
27. Design a data mesh architecture using AWS Lake Formation and AWS Glue. AWS Big Data Blog
28. Data Management Challenges in Production Machine Learning (slides)
29. The Missing Piece of Data Discovery and Observability Platforms: Open Standard for Metadata
30. Automating Data Protection at Scale
31. A curated list of awesome pipeline toolkits

32. Data Mesh Archtitecture
33. The Essential Guide to Data Exploration in Machine Learning (by NimbleBox.ai)

MLOps: Model Deployment and Serving

▼ Click to expand!

1. AI Infrastructure for Everyone: DeterminedAI
2. Deploying R Models with MLflow and Docker
3. What Does it Mean to Deploy a Machine Learning Model?
4. Software Interfaces for Machine Learning Deployment
5. Batch Inference for Machine Learning Deployment
6. AWS Cost Optimization for ML Infrastructure - EC2 spend
7. CI/CD for Machine Learning & AI
8. Itaú Unibanco: How we built a CI/CD Pipeline for machine learning with ***online training*** in Kubeflow
9. 101 For Serving ML Models
10. Deploying Machine Learning models to production — **Inference service architecture patterns**
11. Serverless ML: Deploying Lightweight Models at Scale
12. ML Model Rollout To Production. Part 1 | Part 2
13. Deploying Python ML Models with Flask, Docker and Kubernetes
14. Deploying Python ML Models with Bodywork
15. Framework for a successful Continuous Training Strategy. When should the model be retrained? What data should be used? What should be retrained? A data-driven approach
16. Efficient Machine Learning Inference. The benefits of multi-model serving where latency matters

MLOps: Testing, Monitoring and Maintenance

▼ Click to expand!

1. Building dashboards for operational visibility (AWS)
2. Monitoring Machine Learning Models in Production
3. Effective testing for machine learning systems
4. Unit Testing Data: What is it and how do you do it?
5. How to Test Machine Learning Code and Systems (Accompanying code)
6. Wu, T., Dong, Y., Dong, Z., Singa, A., Chen, X. and Zhang, Y., 2020. Testing Artificial Intelligence System Towards Safety and Robustness: State of the Art. IAENG

International Journal of Computer Science, 47(3).
7. Multi-Armed Bandits and the Stitch Fix Experimentation Platform
8. A/B Testing Machine Learning Models
9. Data validation for machine learning. Polyzotis, N., Zinkevich, M., Roy, S., Breck, E. and Whang, S., 2019. Proceedings of Machine Learning and Systems
10. Testing machine learning based systems: a systematic mapping
11. Explainable Monitoring: Stop flying blind and monitor your AI
12. WhyLogs: Embrace Data Logging Across Your ML Systems
13. Evidently AI. Insights on doing machine learning in production. (Vendor blog.)
14. The definitive guide to comprehensively monitoring your AI
15. Introduction to Unit Testing for Machine Learning
16. Production Machine Learning Monitoring: Outliers, Drift, Explainers & Statistical Performance
17. Test-Driven Development in MLOps Part 1
18. Domain-Specific Machine Learning Monitoring
19. Introducing ML Model Performance Management (Blog by fiddler)
20. What is ML Observability? (Arize AI)
21. Beyond Monitoring: The Rise of Observability (Arize AI & Monte Carlo Data)
22. Model Failure Modes (Arize AI)
23. Quick Start to Data Quality Monitoring for ML (Arize AI)
24. Playbook to Monitoring Model Performance in Production (Arize AI)
25. Robust ML by Property Based Domain Coverage Testing (Blog by Efemarai)
26. Monitoring and explainability of models in production
27. Beyond Monitoring: The Rise of Observability
28. ML Model Monitoring – 9 Tips From the Trenches. (by NU bank)
29. Model health assurance at LinkedIn. By LinkedIn Engineering
30. How to Trust Your Deep Learning Code (Accompanying code)
31. Estimating Performance of Regression Models Without Ground-Truth (Using NannyML)
32. How Hyperparameter Tuning in Machine Learning Works (by NimbleBox.ai)

MLOps: Infrastructure & Tooling

▼ Click to expand!

1. MLOps Infrastructure Stack Canvas
2. Rise of the Canonical Stack in Machine Learning. How a Dominant New Software Stack Will Unlock the Next Generation of Cutting Edge AI Apps
3. AI Infrastructure Alliance. Building the canonical stack for AI/ML
4. Linux Foundation AI Foundation

5. ML Infrastructure Tools for Production | Part 1 — Production ML — The Final Stage of the Model Workflow | Part 2 — Model Deployment and Serving
6. The MLOps Stack Template (by valohai)
7. Navigating the MLOps tooling landscape
8. MLOps.toys curated list of MLOps projects (by Aporia)
9. Comparing Cloud MLOps platforms, From a former AWS SageMaker PM
10. Machine Learning Ecosystem 101 (whitepaper by Arize AI)
11. Selecting your optimal MLOps stack: advantages and challenges. By Intellerts
12. Infrastructure Design for Real-time Machine Learning Inference. The Databricks Blog
13. The 2021 State of AI Infrastructure Survey
14. AI infrastructure Maturity matrix
15. A Curated Collection of the Best Open-source MLOps Tools. By Censius
16. Best MLOps Tools to Manage the ML Lifecycle (by NimbleBox.ai)

MLOps Papers

A list of scientific and industrial papers and resources about Machine Learning operalization since 2015. See more.

Talks About MLOps

▼ Click to expand!

1. "MLOps: Automated Machine Learning" by Emmanuel Raj
2. DeliveryConf 2020. "Continuous Delivery For Machine Learning: Patterns And Pains" by Emily Gorcenski
3. MLOps Conference: Talks from 2019
4. Kubecon 2019: Flyte: Cloud Native Machine Learning and Data Processing Platform
5. Kubecon 2019: Running LargeScale Stateful workloads on Kubernetes at Lyft
6. A CI/CD Framework for Production Machine Learning at Massive Scale (using Jenkins X and Seldon Core)
7. MLOps Virtual Event (Databricks)
8. MLOps NY conference 2019
9. MLOps.community YouTube Channel
10. MLinProduction YouTube Channel
11. Introducing MLflow for End-to-End Machine Learning on Databricks. Spark+AI Summit 2020. Sean Owen
12. MLOps Tutorial #1: Intro to Continuous Integration for ML
13. Machine Learning At Speed: Operationalizing ML For Real-Time Data Streams (2019)
14. Damian Brady - The emerging field of MLops

15. MLOps - Entwurf, Entwicklung, Betrieb (INNOQ Podcast in German)
16. Instrumentation, Observability & Monitoring of Machine Learning Models
17. Efficient ML engineering: Tools and best practices
18. Beyond the jupyter notebook: how to build data science products
19. An introduction to MLOps on Google Cloud (First 19 min are vendor-, language-, and framework-agnostic. @visenger)
20. How ML Breaks: A Decade of Outages for One Large ML Pipeline
21. Clean Machine Learning Code: Practical Software Engineering
22. Machine Learning Engineering: 10 Fundamentale Praktiken
23. Architecture of machine learning systems (3-part series)
24. Machine Learning Design Patterns
25. The laylist that covers techniques and approaches for model deployment on to production
26. ML Observability: A Critical Piece in Ensuring Responsible AI (Arize AI at Re-Work)
27. ML Engineering vs. Data Science (Arize AI Un/Summit)
28. SRE for ML: The First 10 Years and the Next 10
29. Demystifying Machine Learning in Production: Reasoning about a Large-Scale ML Platform
30. Apply Conf 2022
31. Databricks' Data + AI Summit 2022
32. RE•WORK MLOps Summit 2022
33. Annual MLOps World Conference

Existing ML Systems

▼ Click to expand!

1. Introducing FBLearner Flow: Facebook’s AI backbone
2. TFX: A TensorFlow-Based Production-Scale Machine Learning Platform
3. Accelerate your ML and Data workflows to production: Flyte
4. Getting started with Kubeflow Pipelines
5. Meet Michelangelo: Uber’s Machine Learning Platform
6. Meson: Workflow Orchestration for Netflix Recommendations
7. What are Azure Machine Learning pipelines?
8. Uber ATG’s Machine Learning Infrastructure for Self-Driving Vehicles
9. An overview of ML development platforms
10. Snorkel AI: Putting Data First in ML Development
11. A Tour of End-to-End Machine Learning Platforms
12. Introducing WhyLabs, a Leap Forward in AI Reliability
13. Project: Ease.ml (ETH Zürich)

14. Bodywork: model-training and deployment automation
15. Lessons on ML Platforms — from Netflix, DoorDash, Spotify, and more
16. Papers & tech blogs by companies sharing their work on data science & machine learning in production. By Eugen Yan
17. How do different tech companies approach building internal ML platforms? (tweet)
18. Declarative Machine Learning Systems
19. StreamING Machine Learning Models: How ING Adds Fraud Detection Models at Runtime with Apache Flink

Machine Learning

▼ Click to expand!

1. Book, Aurélien Géron,"Hands-On Machine Learning with Scikit-Learn and TensorFlow"
2. Foundations of Machine Learning
3. Best Resources to Learn Machine Learning
4. Awesome TensorFlow
5. "Papers with Code" - Browse the State-of-the-Art in Machine Learning
6. Zhi-Hua Zhou. 2012. Ensemble Methods: Foundations and Algorithms. Chapman & Hall/CRC.
7. Feature Engineering for Machine Learning. Principles and Techniques for Data Scientists. By Alice Zheng, Amanda Casari
8. Google Research: Looking Back at 2019, and Forward to 2020 and Beyond
9. O'Reilly: The road to Software 2.0
10. Machine Learning and Data Science Applications in Industry
11. Deep Learning for Anomaly Detection
12. Federated Learning for Mobile Keyboard Prediction
13. Federated Learning. Building better products with on-device data and privacy on default
14. Federated Learning: Collaborative Machine Learning without Centralized Training Data
15. Yang, Q., Liu, Y., Cheng, Y., Kang, Y., Chen, T. and Yu, H., 2019. Federated learning. Synthesis Lectures on Artificial Intelligence and Machine Learning, 13(3). Chapters 1 and 2.
16. Federated Learning by FastForward
17. THE FEDERATED & DISTRIBUTED MACHINE LEARNING CONFERENCE
18. Federated Learning: Challenges, Methods, and Future Directions
19. Book: Molnar, Christoph. "Interpretable machine learning. A Guide for Making Black Box Models Explainable", 2019
20. Book: Hutter, Frank, Lars Kotthoff, and Joaquin Vanschoren. "Automated Machine Learning". Springer,2019.
21. ML resources by topic, curated by the community.

22. An Introduction to Machine Learning Interpretability, by Patrick Hall, Navdeep Gill, 2nd Edition. O'Reilly 2019
23. Examples of techniques for training interpretable machine learning (ML) models, explaining ML models, and debugging ML models for accuracy, discrimination, and security.
24. Paper: "Machine Learning in Python: Main developments and technology trends in data science, machine learning, and artificial intelligence", by Sebastian Raschka, Joshua Patterson, and Corey Nolet. 2020
25. Distill: Machine Learning Research
26. AtHomeWithAI: Curated Resource List by DeepMind
27. Awesome Data Science
28. Intro to probabilistic programming. A use case using Tensorflow-Probability (TFP)
29. Dive into Snorkel: Weak-Superversion on German Texts. inovex Blog
30. Dive into Deep Learning. An interactive deep learning book with code, math, and discussions. Provides NumPy/MXNet, PyTorch, and TensorFlow implementations
31. Data Science Collected Resources (GitHub repository)
32. Set of illustrated Machine Learning cheatsheets
33. "Machine Learning Bookcamp" by Alexey Grigorev
34. 130 Machine Learning Projects Solved and Explained
35. Machine learning cheat sheet
36. Stateoftheart AI. An open-data and free platform built by the research community to facilitate the collaborative development of AI
37. Online Machine Learning Courses: 2020 Edition
38. End-to-End Machine Learning Library
39. Machine Learning Toolbox (by Amit Chaudhary)
40. Causality for Machine Learning
41. Causal Inference for the Brave and True
42. Causal Inference
43. A resource list for causality in statistics, data science and physics
44. Learning from data. Caltech
45. Machine Learning Glossary
46. Book: "Distributed Machine Learning Patterns". 2022. By Yuan Tang. Manning
47. Machine Learning for Beginners - A Curriculum
48. Making Friends with Machine Learning. By Cassie Kozyrkov
49. Machine Learning Workflow - A Complete Guide (by NimbleBox.ai)
50. Performance Metrics to Monitor in Machine Learning Projects (by NimbleBox.ai)

Software Engineering

▼ Click to expand!

1. The Twelve Factors
2. Book "Accelerate: The Science of Lean Software and DevOps: Building and Scaling High Performing Technology Organizations", 2018 by Nicole Forsgren et.al
3. Book "The DevOps Handbook" by Gene Kim, et al. 2016
4. State of DevOps 2019
5. Clean Code concepts adapted for machine learning and data science.
6. School of SRE
7. 10 Laws of Software Engineering That People Ignore
8. The Patterns of Scalable, Reliable, and Performant Large-Scale Systems
9. The Book of Secret Knowledge
10. SHADES OF CONWAY'S LAW
11. Engineering Practices for Data Scientists

Product Management for ML/AI

▼ Click to expand!

1. What you need to know about product management for AI. A product manager for AI does everything a traditional PM does, and much more.
2. Bringing an AI Product to Market. Previous articles have gone through the basics of AI product management. Here we get to the meat: how do you bring a product to market?
3. The People + AI Guidebook
4. User Needs + Defining Success
5. Building machine learning products: a problem well-defined is a problem half-solved.
6. Talk: Designing Great ML Experiences (Apple)
7. Machine Learning for Product Managers
8. Understanding the Data Landscape and Strategic Play Through Wardley Mapping
9. Techniques for prototyping machine learning systems across products and features
10. Machine Learning and User Experience: A Few Resources
11. AI ideation canvas
12. Ideation in AI
13. 5 Steps for Building Machine Learning Models for Business. By shopify engineering
14. Metric Design for Data Scientists and Business Leaders

The Economics of ML/AI

▼ Click to expand!

1. Book: "Prediction Machines: The Simple Economics of Artificial Intelligence"
2. Book: "The AI Organization" by David Carmona
3. Book: "Succeeding with AI". 2020. By Veljko Krunic. Manning Publications

4. A list of articles about AI and the economy
5. Gartner AI Trends 2019
6. Global AI Survey: AI proves its worth, but few scale impact
7. Getting started with AI? Start here! Everything you need to know to dive into your project
8. 11 questions to ask before starting a successful Machine Learning project
9. What AI still can't do
10. Demystifying AI Part 4: What is an AI Canvas and how do you use it?
11. A Data Science Workflow Canvas to Kickstart Your Projects
12. Is your AI project a nonstarter? Here's a reality check(list) to help you avoid the pain of learning the hard way
13. What is THE main reason most ML projects fail?
14. Designing great data products. The Drivetrain Approach: A four-step process for building data products.
15. The New Business of AI (and How It's Different From Traditional Software)
16. The idea maze for AI startups
17. The Enterprise AI Challenge: Common Misconceptions
18. Misconception 1 (of 5): Enterprise AI Is Primarily About The Technology
19. Misconception 2 (of 5): Automated Machine Learning Will Unlock Enterprise AI
20. Three Principles for Designing ML-Powered Products
21. A Step-by-Step Guide to Machine Learning Problem Framing
22. AI adoption in the enterprise 2020
23. How Adopting MLOps can Help Companies With ML Culture?
24. Weaving AI into Your Organization
25. What to Do When AI Fails
26. Introduction to Machine Learning Problem Framing
27. Structured Approach for Identifying AI Use Cases
28. Book: "Machine Learning for Business" by Doug Hudgeon, Richard Nichol, O'reilly
29. Why Commercial Artificial Intelligence Products Do Not Scale (FemTech)
30. Google Cloud's AI Adoption Framework (White Paper)
31. Data Science Project Management
32. Book: "Competing in the Age of AI" by Marco Iansiti, Karim R. Lakhani. Harvard Business Review Press. 2020
33. The Three Questions about AI that Startups Need to Ask. The first is: Are you sure you need AI?
34. Taming the Tail: Adventures in Improving AI Economics
35. Managing the Risks of Adopting AI Engineering
36. Get rid of AI Saviorism
37. Collection of articles listing reasons why data science projects fail

38. How to Choose Your First AI Project by Andrew Ng
39. How to Set AI Goals
40. Expanding AI's Impact With Organizational Learning
41. Potemkin Data Science
42. When Should You Not Invest in AI?
43. Why 90% of machine learning models never hit the market. Most companies lack leadership support, effective communication between teams, and accessible data

Model Governance, Ethics, Responsible AI

This topic is extracted into our new Awesome ML Model Governace repository

MLOps: People & Processes

▼ Click to expand!

1. Scaling An ML Team (0–10 People)
2. The Knowledge Repo project is focused on facilitating the sharing of knowledge between data scientists and other technical roles.
3. Scaling Knowledge at Airbnb
4. Models for integrating data science teams within companies A comparative analysis
5. How to Write Better with The Why, What, How Framework. How to write design documents for data science/machine learning projects? (by Eugene Yan)
6. Technical Writing Courses
7. Building a data team at a mid-stage startup: a short story. By Erik Bernhardsson
8. The Cultural Benefits of Artificial Intelligence in the Enterprise. by Sam Ransbotham, François Candelon, David Kiron, Burt LaFountain, and Shervin Khodabandeh

Newsletters About MLOps, Machine Learning, Data Science and Co.

▼ Click to expand!

1. ML in Production newsletter
2. MLOps.community
3. Andriy Burkov newsletter
4. Decision Intelligence by Cassie Kozyrkov
5. Laszlo's Newsletter about Data Science
6. Data Elixir newsletter for a weekly dose of the top data science picks from around the web. Covering machine learning, data visualization, analytics, and strategy.
7. The Data Science Roundup by Tristan Handy

8. Vicki Boykis Newsletter about Data Science
9. KDnuggets News
10. Analytics Vidhya, Any questions on business analytics, data science, big data, data visualizations tools and techniques
11. Data Science Weekly Newsletter: A free weekly newsletter featuring curated news, articles and jobs related to Data Science
12. The Machine Learning Engineer Newsletter
13. Gradient Flow helps you stay ahead of the latest technology trends and tools with in-depth coverage, analysis and insights. See the latest on data, technology and business, with a focus on machine learning and AI
14. Your guide to AI by Nathan Benaich. Monthly analysis of AI technology, geopolitics, research, and startups.
15. O'Reilly Data & AI Newsletter
16. deeplearning.ai’s newsletter by Andrew Ng
17. Deep Learning Weekly
18. Import AI is a weekly newsletter about artificial intelligence, read by more than ten thousand experts. By Jack Clark.
19. AI Ethics Weekly
20. Announcing Projects To Know, a weekly machine intelligence and data science newsletter
21. TWIML: This Week in Machine Learning and AI newsletter
22. featurestore.org: Monthly Newsletter on Feature Stores for ML
23. DataTalks.Club Community: Slack, Newsletter, Podcast, Weeekly Events
24. Machine Learning Ops Roundup
25. Data Science Programming Newsletter by Eric Ma
26. Marginally Interesting by Mikio L. Braun
27. Synced
28. The Ground Truth: Newsletter for Computer Vision Practitioners
29. SwirlAI: Data Engineering, MLOps and overall Data focused Newsletter by Aurimas Griciūnas