МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
Одеський національний технологічний університет
Університет Інформатики і прикладних знань, м.Лодзь, Польща
Національний технічний університет України «Київський політехнічний інститут»
Навчально-науковий інститут комп'ютерних систем і технологій «Індустрія 4.0» ім. П.М. Платонова

XXIII Всеукраїнська науково-технічна конференція молодих вчених, аспірантів та студентів

«СТАН, ДОСЯГНЕННЯ ТА ПЕРСПЕКТИВИ ІНФОРМАЦІЙНИХ СИСТЕМ І ТЕХНОЛОГІЙ»

Матеріали конференції

Одеса

20-21 квітня 2023 р.

Стан, досягнення та перспективи інформаційних систем і технологій / Матеріали XXIII Всеукраїнської науково-технічної конференції молодих вчених, аспірантів та студентів. Одеса, 20-21 квітня 2023 р. - Одеса, Видавництво ОНТУ, 2023 р. – 449 с.

Збірник включає матеріали доповідей учасників конференції, які об'єднані за тематичними напрямками конференції.

Збірник буде корисним як для фахівців і працівників фірм, зайнятих в області IT, так і для викладачів, магістрів і студентів вищих навчальних закладів, які навчаються за напрямами і спеціальностями програмного забезпечення, обчислювальної техніки і автоматизованих систем, прикладної математики та обробки інформації, буде корисним професіоналам з комп'ютерного моделювання та розробки комп'ютерних ігор.

Результати досліджень у збірнику представляють собою своєрідний зріз сучасного стану справ в перерахованих галузях знань, який може допомогти як фахівцям, так і студентам університетів скласти загальну картину розвитку інформаційних технологій та пов'язаних з ними питань.

Наукові праці згруповані за напрямками роботи конференції та наведені в алфавітному порядку прізвищ авторів.

Матеріали (тези доповідей) друкуються в авторській редакції. Відповідальність за якість та зміст публікацій несе автор.

Матеріали подано українською та англійською мовами.

Редактор збірника Котлик С.В.

3. The beginner's guide to semantic search: Examples and tools, Енн Смарті, Грудень 2019, https://www.searchenginewatch.com/2019/12/16/the-beginners-guide-to-semantic-search/
4. О.Р. Смиш, О.В. Золотаревич, "NLP: створення парсера для структурування тексту вакансій", 2022
5. UDPipe trainable pipeline, Czech Republic http://ufal.mff.cuni.cz/udpipe. UDPipe web service is available on http(s)://lindat.mff.cuni.cz/services/udpipe/api/.

MLOPS AS AN APPROACH TO MANAGE MACHINE LEARNING MODELS LIFECYCLES

KOLOMYTSEV A.M. (a.kolomytsev@khai.edu),
KUZNETSOVA Y.A. (y.kuznetsova@khai.edu), **SHULGA I.M.** (i.shulga@khai.edu)
National Aerospace University "Kharkiv Aviation Institute"

This report explores the evolution of software development and delivery and the emergence of MLOps - a set of practices that bridged the gap between development and data analytics teams. DevOps introduced automation, continuous integration, and continuous delivery, enabling organizations to deliver high-quality software at scale and with greater speed. MLOps has adapted the best practices of DevOps and SRE to maintain and deploy ML models with high reliability and efficiency. MLOps allows for automation, observability through metrics, and the ability for data scientists to deploy their models without manual intervention from IT operations teams.

Over the last few decades, we've seen a dramatic shift in the way organizations approach software development and delivery. In the early days of computing, IT operations teams were responsible for managing large mainframe systems and keeping them running smoothly. As technology evolved and distributed systems became the norm, IT operations teams had to adapt to new challenges and find ways to manage complex infrastructure at scale [1].

The rise of DevOps movement in the early 2000s marked a turning point in the world of software development and IT operations. Bass, Weber and Zhu defined it "as a set of practices intended to reduce the time between committing a change to a system and the change being placed into normal production,while ensuring high quality" [2]. DevOps introduced a set of practices that broke down the traditional silos between development and operations teams, enabling them to work together more collaboratively and efficiently. DevOps brought automation, continuous integration, and continuous delivery to the forefront of software development, enabling organizations to deliver high-quality software at scale and with greater speed. There are multiple approaches to DevOps, one of the most famous is SRE, developed and used by Google: SRE emphasizes the importance of service-level objectives - a measurable target for the level of service a system should provide to its users. SRE is primarily focused on the reliability and operability of services and tries to cultivate a blameless culture where failure is treated as an opportunity to learn and improve [3].

This transition that happened to regular system software deployment also happened to ML models. Before the emergence of MLOps, the deployment and management of machine learning (ML) models were often ad hoc and unstructured. ML models were typically developed by data scientists in silos and then handed over to IT or operations teams to deploy in production environments. This handoff was often a complex and error-prone process, requiring significant manual intervention and coordination. Only a couple of big tech giants were able to successfully automate this [4].

Subramanya defines MLOps as "a set of practices that aims to maintain anddeploy Machine Learning code and models with high reliability and efficiency"[5], which is really close to definition of the DevOps given by Bass, Weber, Zhu. Basically, IT operations engineers took practices they learned while deploying regular system software and applied this to ML models and it worked. MLOps, following DevOps and SRE guidelines is focused on automation, observability through metrics and giving people responsible for the AI models developlemnt tools to deploy their models without manual intervention from Ops team. MLOps allows us to use familiar Blue/Green deploymeny processes and A/B testing, we used for regular applications, to manage our ML pipelines.

One of the key challenges that MLOps aims to address is the versioning and reproducibility of ML models. This is critical in ensuring that models are accurately and consistently deployed across different environments, without causing any unintended consequences. MLOps also emphasizes the importance of monitoring and maintaining the performance of ML models in production, to ensure that they continue to provide accurate predictions over time.

MLOps tools and platforms have emerged to help streamline the entire ML lifecycle, from data preparation and model training to deployment and monitoring. These tools enable data scientists to focus on building and improving models, while providing IT and operations teams with the necessary infrastructure to deploy and manage ML models at scale.

In conclusion: DevOps practices changed software development processes for the better, allowing us to have better automation, observability and helping developers to know more about their production environment. By applying DevOps principles and practices to ML development and deployment, MLOps aims to make the process of managing ML models in production environments more efficient, reliable, and scalable. And, seeing how widespread DevOps, as set of practices become for software development, we can be sure the same would happen with MLOps in ML/AI development teams.

MLOps is still in its early stages, but it's quickly gaining popularity among organizations that are investing heavily in machine learning and artificial intelligence. As the demand for ML models continues to grow, MLOps will become increasingly important for organizations that want to deploy and manage their models at scale.

List of references

1. M. Gokarna and R. Singh, "DevOps: A Historical Review and Future Works," in 2021 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS), Feb. 2021, pp. 366–371. doi: 10.1109/ICCCIS51004.2021.9397235.

2. L. Bass, I. Weber, and L. Zhu, DevOps: A Software Architect's Perspective. Addison-Wesley Professional, 2015.

3. B. Beyer, C. Jones, J. Petoff, and N. R. Murphy, Site Reliability Engineering: How Google Runs Production Systems. O'Reilly Media, Incorporated, 2016.

4. J. Lin and A. Kolcz, "Large-scale machine learning at twitter," in Proceedings of the 2012 ACM SIGMOD International Conference on Management of Data, New York, NY, USA, May 2012, pp. 793–804. doi: 10.1145/2213836.2213958.

5. R. Subramanya, S. Sierla, and V. Vyatkin, "From DevOps to MLOps: Overview and Application to Electricity Market Forecasting," Applied Sciences, vol. 12, no. 19, Art. no. 19, Jan. 2022, doi: 10.3390/app12199851.