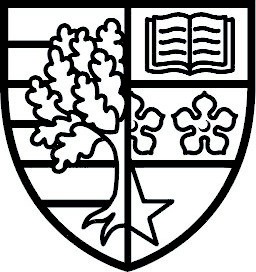
MLOps-Driven Music Recommendation System

Usman Jamshaid Khan

BSc (Hons.) Computer Science Honours Dissertation

*Supervised by* Dr Abrar Ullah

Deliverable 1: Final Year Dissertation



Heriot-Watt University

School of Mathematical and Computer Sciences

BSc Computer Science (Software Engineering)

September 2024

The copyright in this dissertation is owned by the author. Any quotation from the dissertation or use of any of the information contained in it must be ac- knowledged as the source of the quotation or information.

DECLARATION

I, Usman Jamshaid Khan, confirm that this work submitted for assessment is my own and is ex- pressed in my own words. Any uses made within it of the works of other authors in any form (e.g., ideas, equations, figures, text, tables, programs) are properly acknowledged at any point of their use. A list of the references employed is included.

Signed: Usman

Date: 14/10/24

i

ABSTRACT

The goal of Machine Learning Operations (MLOps) is to build and implement a program to have quick production. It has gained recognition as a critical answer to the problems presented in the Machine learning (ML) sector. Like DevOps, MLOps uses development, testing, and monitoring to enhance ML production. MLOps comprises procedures for automation, continuous integration and continuous development (CI/CD), and model governance to guarantee that ML models can be swiftly and reliably deployed to production. These days, ML projects executed without MLOps may result in subpar management and sluggish deployment cycles. Development may be sped up in the long run by using MLOps to address these problems using practices, ideas, and development.

With the help of MLOps, this project develops a music recommendation system that makes song recommendations based on face expression detection. The recommendation system will make use of machine learning techniques like Convolutional Neural Networks (CNNs), which are essentially capable of identifying human emotions and recommending music in line with those feelings. The system may achieve efficiency, maintainability, and risk reduction by incorporating MLOps, ensuring that the deployment is ongoing and under observation. It can facilitate automation for model deployment, training, and ongoing monitoring, all of which will eventually result in problems-free improvements. In the end, this project demonstrates the effectiveness of using facial recognition to offer personalized, emotion-matched music recommendations in real time, as well as the resilience of MLOps in managing the complex demands of machine learning.

iii

Table of Contents

### [Declaration](#_bookmark0) i

[Abstract](#_bookmark1) iii

[Table of Contents](#_bookmark3) vii

[List of Figures](#_bookmark4) ix

[List of Tables](#_bookmark5) xi

1. [Introduction](#_bookmark6) 1
   1. [Motivation](#_bookmark7) 1
   2. [Aim and Objectives](#_bookmark8) 1
   3. [Contributions](#_bookmark9) 2
   4. [Organization](#_bookmark10) 2
2. [Background](#_bookmark11) 3
   1. Machine Learning Development Challenges 3
   2. DevOps 4
   3. MLOps 5
   4. MLOps Challenges 6
   5. Recommendation System as a Use Case 7
   6. Music Recommendation System Using Facial Recognition 8
   7. MLOps in Music Recommendation System 9
   8. Related Work 9
   9. Summary 9
3. [Method](#_bookmark15) 10
   1. [Back-end](#_bookmark16) 5
   2. [Front-end](#_bookmark22) 6
   3. [Summary](#_bookmark26) 7
4. [Requirements](#_bookmark27) 9
   1. [Results about B](#_bookmark28) 9
   2. [Results about A](#_bookmark29) 9
   3. [Summary](#_bookmark30) 9
5. [Evaluation](#_bookmark38) 15
   1. [Motivation and Goals](#_bookmark39) 15
   2. [Contributions](#_bookmark40) 15
   3. [Limitations and Future Work](#_bookmark41) 15

[References](#_bookmark42) 17

A Appendix: Project Management

List of Figures

1. [Front-end mockup](#_bookmark23) 6
2. [Flowcharts](#_bookmark25) 7
3. [Mockup of front-end.](#_bookmark48) 21

ix

List of Tables

1. [Table of some properties](#_bookmark35)

# INTRODUCTION

Machine learning has greatly improved many systems and one of them is the recommendation systems which can be related to music, e-commerce, movies, etc. These systems are advanced, but the deployment and maintenance of these systems remains challenging because of the complexity of maintaining the scalability and efficiency of models. This is where MLOps emerges, which is a DevOps concept modified for the management and optimization of machine learning processes. In order to ensure that machine learning models operate dependably and efficiently in production settings, MLOps optimizes the whole machine learning lifecycle, from data preprocessing and model development to deploying, and monitoring. This dissertation’s main goal is to create a music recommendation system that uses face recognition to determine the user’s moods and make appropriate song suggestions. The project intends to guarantee effective deployment and maintenance of the face emotion recognition-based music recommendation system by implementing MLOps techniques. Using the music recommendation system as a use case, this research will investigate how MLOps concepts may enable the continuous delivery, oversight, and scalability of recommendation system that uses machine learning.

## Motivation

Machine learning models have shown a lot of progress recently for resolving challenging issues in a variety of industries. Nevertheless, regardless of its potential, ML model implementation and maintenance continue to be extremely difficult. Delivering ML products that are scalable and reliable might be challenging due to difficulties caused by separate techniques from data scientists and ml engineers. By incorporating DevOps ideas into the ML process, MLOps fills the gap by facilitating continuous integration, continuous delivery, training, and monitoring and ensures that models are effectively maintained. In this dissertation, MLOps will be implemented to show how it improves the efficacy of a face emotion recognition-based music recommendation system. As models are developed on face emotions to suggest music, they should constantly adjust to new datasets and enhance the suggestions. By integrating MLOps the system may be designed, implemented, and monitored in an automatic and flexible way.

## Aim and Objectives

In this dissertation, the main aim is to investigate and apply MLOps approaches to create a machine learning pipeline for a music recommendation system that is reliable and scalable. The project will involve investigation into existing MLOps practices and their influence on machine learning processes and will be succeeded by the hands-on application of these methods. From implementation to deployment, and continuous procedures, the project will show how MLOps can greatly improve the operation of machine learning models workflow.

The main objectives are:

* Conduct research on MLOps techniques to help with the development of a reliable and expandable machine learning pipeline.
* Develop ML pipelines with MLOps for effective data collection, training of data, and delivery.
* Implement CI/CD pipeline to automate evaluation, verification, delivery and continuous monitoring to identify model drift and automate retraining.
* Develop a face emotion-based music recommendation system using MLOps practices.
* Evaluate the MLOps pipelines for automation, and efficiency.

## Contributions

The contributions of this project are:

* + 1. Adapting MLOps methods to meet the demands of machine learning systems, such as recommendation systems.
    2. Evaluation of how MLOps procedures affect the scalability and reliability of the system.
    3. Implementation of music recommendation system using facial recognition with MLOps procedures.

## Organization

Here is how this dissertation is organized. Following the introduction, motivation, and contribution, Section 2 presents a Background that covers the state-of-the-art research in MLOps, DevOps, and music recommendation systems. Section 3 presents the Methodologies and Requirements where the research methods and requirements analysis will be suggested. Section 4 will display the system architecture and Section 5 will show the evaluation methods that will be utilized to focus on the MLOps pipeline efficiency.

# BACKGROUND

This background discusses recent studies done on the importance of Machine Learning Operations or (MLOps), focusing on why MLOps is a crucial framework in the field of machine learning and its relevance in music recommendation system using facial recognition. The emphasis is on comprehending how MLOps can improve Machine learning systems, recognizing the difficulties, and investigating relevant research in the field.

## Machine Learning Development Challenges

The field of machine learning (ML) enables machines to independently learn from data and previous events in order to recognize patterns in data, categorize data, and forecast outcomes with little assistance from humans (Mitchell & Jordan, 2015). Previously, the machine learning models were typically few in number, making them easier to oversee, or there was lesser organizational interest in comprehending these models and their relationships but with the emergence of decision automation, where choices are made more frequently without human input, machine learning models have become increasingly vital and the management of risk related to these models have become more significant at elevated levels withing the organization (Treveil et al., 2021).

It has always been difficult to integrate machine leaning (ML) models into production, with data scientists, machine learning engineers and other groups encountering various obstacles during the development of production ready models, which results in very small percentage of ML projects even getting to the stage of production and this is why several tools have emerged to enhance processes such as creating models, processing data, training, all aimed to address those challenges and advance in this field (Symeonidis et al., 2022). Figure 1 illustrates the workflow of machine learning life cycle.

A diagram of a process

Description automatically generated

Figure 1: The Machine Learning Life Cycle (Testi et al., 2022)

Three key factors make managing machine learning life cycles challenging. First there are many dependencies since business demands and data are ever-changing, necessitating ongoing alignment between the model’s performance and its initial objectives. Second, because business, data science, and IT teams employ distinct technologies and expertise, communication gaps develop between them. Finally, juggling several responsibilities strain data scientists, who are frequently not software engineers, and this is particularly true when the number of models increases, and they are expected to manage models that they did not initially develop (Treveil et al., 2021).

## DevOps

Prior to the development of Machine Learning Operations principles, businesses faced considerable difficulties in implementing solutions developed with the newest machine learning technology because of the substantial resources needed, which is why MLOps was so important (Alla & Adari, 2020). To understand MLOps we need to comprehend the definition of Development Operations (DevOps). DevOps is a collection of procedures that blends the work procedures of the operations groups with developers of software to provide a shared set of procedures that acts as a cross for both positions (Alla & Adari, 2020). A graph depicting the DevOps workflow is shown in Figure 2. It’s a change in structure where, rather than dispersed groups handling tasks independently, collaborative teams focus on ongoing operational feature deployments (Ebert et al., 2016).

With the implementation of DevOps, software development cycles are accelerated, guaranteeing continued software supply, and lowering overall expenses through reduced service costs-achieved through improved process efficiency in maintaining applications (Alla & Adari, 2020). DevOps tools like continuous integration (CI), continuous delivery (CD), automated testing, and monitoring enable those advantages. By using a continuous integration approach, software development companies aim to incorporate code created by developers and make frequent improvements to it (Deza & Gift, 2021). Continuous delivery is a practice of developing and testing code constantly without human involvement. Among the crucial techniques the most practiced are continuous integration and continuous delivery. Automated testing guarantees that modifications to the code do not create any new bugs or errors (Liu et al., 2023).

Monitoring allows teams to observe and evaluate the system's performance, enabling them to identify any problems. According to Kreuzberger et al. (2023), implementing DevOps techniques may enhance software quality and drastically cut time-to-market by encouraging an increasingly reactive and agile development process.

Diagram of software development process

Description automatically generated

Figure 2: An illustration of a DevOps environments cycle (Alla & Adari, 2020).

## MLOps

Since DevOps has been successfully implemented, businesses are searching for continuous methods for creating machine learning systems (John, Olsson and Bosch, 2021). To handle the difficulties of implementing and maintaining ML models, MLOps applies that concept of DevOps and the procedures of Data Engineering to the ML lifecycle as shown in Fig 3. Model creation, deployment, monitoring and maintenance are all included in MLOps which aims to optimize ML processes and guarantee that models are flexible, repeatable and maintained (Singla, 2023). A fundamental element of MLOps is the implementation of continuous integration and continuous deployment for machine learning models, which streamlines the training, validation and deployment processes for these models, and it allows models to be enhanced progressively without the need for human involvement (Kreuzberger et al., 2023). MLOps brings forth a new approach, alongside continuous integration and continuous deployment, known as continuous training (CT), which focuses on the automatic retraining of models when necessary. Continuous monitoring is also essential as it involves the ongoing evaluation of model performance to identify and resolve issues like model drift or poor data quality (Hymel et al., 2022). Version control is also crucial, which involves continuously developing different stages of machine learning models in order to keep track and revert to earlier versions as necessary. MLOps automates the process of machine learning models, extending beyond traditional software processes and it also combines the workflows of data scientists and Data engineers with those of the developers to enable the continuous delivery of high-quality machine learning models (Alla & Adari, 2020). Figure 4 shows a typical MLOps workflow.

A diagram of machine learning

Description automatically generated

Figure 3: MLOps involves DevOps, ML, and Data Engineering (Hewage and Meedeniya,2022)

A diagram of a model development

Description automatically generated

Figure 4: The MLOps Workflow (Testi et al., 2022)

## MLOps Challenges

MLOps offers advantages, but it also encounters various difficulties. Technical problems like versioning models might be difficult because the models must be reproducible. Since models can act differently in development and production contexts, it can be difficult to ensure reliable performance across a variety of scenarios (Singla, 2023). This is in accordance with the difficulties mentioned by Ruf et al. (2021), and (Ghantous and Gill, 2017), which emphasizes how crucial it is to choose the appropriate tools for MLOps pipelines in order to guarantee quality and conformity with DevOps procedures. Implementing MLOps is further complicated by the intricacy of the deployment of machine learning systems numerous clouds settings (Banerjee et al., 2020). Implementing effective version control system and containerization methods can make the models reproducible and the deployment consistent (Singla, 2023).

The Organization difficulties include promoting teamwork among cross-disciplinary groups, like data scientists, ML engineers, and operations personnel, who might prioritize different objectives and operate with varying processes and diverse technologies and frameworks are frequently used by organizations, which can cause integrity issues and inefficiency (Singla, 2023). Granlund et al. (2021), discusses the difficulties in implementing hybrid models in various businesses and offers strategies to make deploying easier in these situations. The creation of specialized MLOps teams that prioritize interaction and information exchange can promote cross-functional cooperation. Organizational silos may be reduced, and processes streamlined by incorporating MLOps into current DevOps procedures.

Because implementing MLOps methods requires a change in processes and thinking, cultural difficulties include a reluctance to change, and the lack of skills is also a problem because teams must train and develop in order to use MLOps procedures effectively (Singla, 2023). Moreover, it is essential for everyone involved to have a common grasp of ML concepts to ensure effective implementation and training and skill development initiatives can help close gaps in skill while fostering a culture of constant progress and flexibility. (Singla, 2023). Furthermore, as suggested (Hewage and Meedeniya, 2022), the accessibility of technologies that are open source might improve interoperability and facilitate incorporation ain many environments.

Overall, with implementing ML-specific automated testing and continuous integration and continuous deployment the dependability of ML systems will be increased, and the standards and compatibility are enhanced by open-source tools, which make integration into many ecosystems easier (Singla, 2023).

## Recommendation System as a Use Case

## Recommendation System

The necessity for real-time adaptation to users changing preferences, which traditional models cannot provide, makes recommendation systems an ideal application for MLOps. Every time new data is produced, traditional recommendation systems must be retrained, which reduces efficiency and slows the improvements. On the other hand, MLOps automates the deploying and retraining of models, allowing for continuous procedures and quick response to changes. This feature is essential for recommendation systems in customized domains like music, where rapid response of appropriate recommendations improves user engagement. Given the vast number of products accessible online there is a demand for systems that help users find their desired items and recommendation system is a key service that assists users in overcoming the challenges of excessive information (Chen & Chen, 2005).

A recommendation system is a software application that provides strategies to propose recommendations to a user based on the likelihood of their preference. Collaborative filtering, content-based filtering, and hybrid systems are three of the prevalent methods utilized in recommendation systems. Collaborative filtering is based on the actions of user’s similar tastes to generate recommendations. Content-based filtering focuses on the characteristics of items to suggest similar options. In collaborative filtering, if a user prefers X, and Y is identical to X, eventually the user might also prefer Y and in content-based filtering, if a user X and a user Y are identical, then user X may prefer what user Y prefers (Fessahaye et al.) Hybrid systems merge both strategies to improve precision and resilience. Strong machine learning models that can manage massive data sets, update instantly and adjust to shifting consumer suggestions are necessary for recommendation systems to function well. The application of cutting-edge ML methods, such as deep learning and reinforcement learning has significantly enhanced the effectiveness of recommendation system making them more precise and effective (Wang et al., 2023).

## Music Recommendation System using Facial Recognition

The Music Recommendation system that uses facial recognition can provide incredibly customized experiences by recognizing user emotions and pairing music appropriately. This system examines the user's facial expressions via a camera and employs machine learning algorithms to understand the emotions being shown and recommend music based on the users’ mood (Pradhan et al., 2022). According to Mishra et al. (2024), the system assesses the user’s mood to provide a playlist that corresponds with the way they are feeling. For example, it may select lively songs from a playlist to change a depressed mood or offer melodies to enhance positive thoughts when the user’s mood is good. Emotion detection using facial recognition can be achieved by deep learning techniques. Deep Leaning, utilizing artificial neural networks that feature numerous hidden layers to replicate the human brains cerebral cortex has greatly progressed the field of computer vision, especially via convolutional neural network (CNN), a robust algorithm, effectively handles millions of parameters by employing filters to analyze 2D images and generate data (Chauhan et al., 2018). A CNN model for face detection analyzes face images to identify features, such as facial landmarks, and works with a dataset containing emotions and this dataset goes through preprocessing steps that include reshaping, resizing and converting to arrays and the CNN then interprets the pixel information to determine the user’s emotions (Shakti et al., 2023). When the emotion is detected, a song will be played that matches the emotion.

## MLOps Implementation in Music Recommendation System

The implementation of a music recommendation system using MLOps practices is crucial as it will manage the life cycle of a model. MLOps automates machine learning operations and improves dependability to expedite the development and deployment of a music recommendation system. According to Susila (2022), MLOps creates a systematic pipeline for model creation, deployment and monitoring by including continuous integration and continuous delivery (CI/CD) into the systems model lifecycle, drawing inspiration from DevOps. The report further explains that the systems speed and precision in suggesting music according to identified emotions are improved by this method, which guarantees that every update of model, trained on face recognition data and emotion signals, is effectively evaluated and implemented. In order to maintain an efficient and flexible recommendation system that keeps up with changing user information MLOps automates the models process, which makes it more scalable and permits immediate changes, and facilities the continuous monitoring. All things considered, MLOps will offer crucial protocols for setting up and managing music recommendation systems that are always updated effortlessly.

## Related Work

There have been numerous studies where MLOps has been implemented in systems to enhance development and deployment. In the study of Susila (2022), MLOps was used to implement the music recommendation system in order to avoid the problem of manually training data every time there’s update in data, which is time consuming and inefficient. They ensured that model changes were easily incorporated by automating the retraining process with MLOps. Furthermore, they mentioned that the stages are separated into Continuous Integration and Continuous Delivery. Continuous integration involved model parameters being regularly assessed to get higher precision and the model was only permitted to move to subsequent phase after it satisfied the necessity accuracy requirements. The verified model was then deployed to systems during the continuous delivery stage so that it can be accessible.

In the area of electricity market forecasting, Subramanya et al. (2022) emphasize how MLOps tackles the operational difficulties encountered with machine learning models in production. The difficulties in the field include model drift, which occurs when models become less accurate over time due to changing market conditions and the challenges associated with the frequent retraining and deployment of models that can lead to delays. It was further elaborated that since DevOps doesn’t address the complexities of Machine learning cycles MLOps is proposed as a solution which automates the workflow of machine learning models. Subramanya et al. (2022) suggests an MLOps pipeline that is appropriate to the electricity market and incorporate technologies for continuous integration and deployment that will allow for quicker responsiveness to updates while preserving model correctness.

## Summary

In summary, MLOps contributes to addressing major issues in ML development and stressing how crucial it is to simplify the development and monitoring of ML models. With an emphasis on face recognition enhanced music recommendation system, it explores how MLOps may streamline processes for systems that tailor user experiences according to emotions. These systems are made more adaptable, reliable, and flexible by the procedures of MLOps. MLOps is crucial for contemporary ML systems since it greatly fortifies the structure required for recommendation systems based on emotions.

# METHODOLOGY

This methodology section will cover the integration of machine learning operations (MLOps) methods to expedite the development, deployment, testing and monitoring of machine learning models. This study includes the application of MLOps to a face emotion-based music recommendation system. The main aim is to demonstrate how well MLOps manages the whole lifespan of machine learning models. The methodology is organized as follows: Initially how MLOps pipelines are designed will be discussed, with an emphasis on technologies that make model procedures easier, such as MLflow and Jenkins. Next, there will be discussion on how these pipelines will work with the implementation of face emotion-based music recommendation system and finally the justification and ethical issues will be covered. Python will be the main programming language utilized for the duration of this project because of its extensive library of machine learning tools, which makes it the perfect choice.

## 3.1 Theoretical basis

MLOps, a collection of procedures designed to streamline the construction and operation of machine learning systems, is the theoretical basis of this study. Through the automation of processes like deployment, versioning, testing, and monitoring, MLOps enhances the scalability, reproducibility, and effectiveness of machine learning workflows.

## 3.2 MLOps Pipelines

Designing and implementing reliable MLOps pipelines to automate crucial ML workflow activities is the core of this research and a range of tools will be utilized when setting up these pipelines. The whole CI/CD procedure will be automated by Jenkins, guaranteeing that models are validated, verified and then deployed automatically with little assistance from humans. Performance measurements will be recorded, model versions will be managed, and experiments will be tracked using MLflow. Git will make it easier to track versions, guaranteeing integrity throughout projects’ many phases. Lastly the system will be containerized by docker, guaranteeing that each element may be delivered in a unform setting during the all the phases. Collectively, these technologies will automate the phases of model development such as model verification, delivery, training, and monitoring, allowing for ongoing upgrades.

Automating the deployment procedure will be made possible via CI/CD pipelines and if modifications are uploaded, these pipelines will start immediately. There will be phases in the pipeline for:

* Model Training: The system will train the model automatically.
* Model Validation: The model will be evaluated whenever there’s a fresh model.
* Testing: Unit and Implementation tests will be conducted for system accuracy and reliability.
* Docker: To guarantee a consistent delivery the whole system will be containerized.
* Deploy: When the model is ready it will be published seamlessly without human intervention.

## 3.3 Model for Face emotion recognition

To detect emotions, a convolutional Neural Network (CNN) will be trained using the FER2013 dataset and for image categorization applications like face emotion recognition, CNNs are especially appropriate. PyTorch will be used for training the model and the images will be preprocessed to improve the model’s precision. The model will be automatically tracked, versioned, and tested once it has been trained and incorporated with the MLOps pipelines and this guarantees that the emotion detection model is continuously updated and operating at its best.

## 3.4 DEAM dataset

The music recommendation algorithm will use the emotion taken from face emotions and songs from DEAM dataset will be compared to the emotions by the algorithm. Music songs in the DEAM dataset have emotional qualities like happy, sad, excitement, fear, anger, and calm and the system will offer tailored music suggestions that correspond with the user’s facial emotion by comparing the emotion characteristics of the song with the emotion identified.

## 3.5 Monitoring

The model’s accuracy will be monitored using MLflow, which will record and display model accuracy and measure the way the model reacts to incoming data and can also allow assessment of multiple models. This method guarantees that models stay updated, and any notable modifications may be promptly found and fixed.

## 3.6 Justification and Ethical Standards

The growing demand for reliable and continuous machine learning processes in production setting is what led to the decision to make MLOps the main approach. Continuous development, continuous training, versioning and monitoring are examples of MLOps methods that offer effective and adaptable approaches to managing the lifespan of machine learning models. This study shows how MLOps pipelines may enhance machine learning systems by using technologies like Jenkins, MLflow, and docker. The Deam dataset and CNNs models are ideally matched to the study’s objectives and offer a strong basis for showcasing the advantages of MLOps workflows. Both FER2013 dataset and Deam dataset are publicly available to use and have anonymous data.

## 3.7 Strengths and Limitations

One of MLOps advantages is its capacity to automate crucial processes like training, testing, and deploying models, which greatly boosts productivity and lower mistakes made by humans. In addition, it also provides scalable solutions, which makes managing datasets simpler and provides a uniform platform for team cooperation. The complicated nature of establishing and managing pipelines is one of MLOps drawbacks, too, and it may be difficult, particularly for groups with little expertise and it may also be challenging to integrate disparate technologies and handling data throughout the pipeline’s phases can be complicated.

## 3.8 Conclusion

In conclusion, the methodology offers an extensive structure for maintaining machine learning models at every stage of their lifespan, and the methodology also shows the development and deployment of pipelines that are automatic using platforms like Jenkins, MLflow, docker, etc. The combination of music recommendation and emotion detection algorithms demonstrates how MLOps may improve practical machine learning systems.

# Requirement Analysis

This part describes the functional requirements (FR) and non-functional requirements (NFR) and includes use case scenario, use case descriptions and a MoSCoW analysis to rank the requirements.

# Use Case Scenario

The system uses a facial expression to determine the emotions and then suggests song that reflects the emotion, and the machine learning model will be trained, tested, deployed and monitored automatically using MLOps workflow.

## 4.2 Functional Requirements

|  |  |  |
| --- | --- | --- |
| ID | Requirement | Priority |
| FR1 | The system must automatically train the emotion recognition model | Must have |
| FR2 | The system must automatically test the emotion recognition model | Must have |
| FR3 | The system must enable CI/CD pipelines for updating and deploying new models | Must have |
| FR4 | The system should continuously monitor the performance of models | Could have |
| FR5 | The system should manage versions of model for tracking | Should have |
| FR6 | The system must be able to retrain when new data is available | Must have |

## Non-functional Requirement

|  |  |  |
| --- | --- | --- |
| ID | Requirement | Priority |
| NFR1 | The system must support accurate model retraining | Must have |
| NFR2 | Models should be reproducible through MLflow monitoring | Could have |
| NFR3 | The system should be scalable to handle big models or datasets. | Could have |

# EVALUATION

The MLOps pipeline evaluation technique is centered on evaluating the machine learning workflows overall performance and efficiency. This also covers continuous model training, deployment, versioning and monitoring. The objective is to guarantee the pipelines smooth operation, effective, and continuous deployment.

# Efficiency of automation

With little human involvement, the MLOPs pipeline should automate the crucial steps of model training, deployment, and testing. The time required to continuously initiate and carry out each procedure, including training and deploying models when fresh data becomes available, will be measured in order to assess the efficiency of automation. Important metrics to assess are deployment speed, duration of testing, training, and pipeline response time to new data.

# Model Versioning

For machine learning models to be reproducible, model versioning is essential. The system’s capacity to monitor various model versions and datasets will be evaluated. A good version control can quickly roll back to earlier model version and the metric for the accuracy of version control includes the success rate formula.

Success Rate = (Number of Success/Total number of Attempts) x 100

# Pipeline Efficiency

The whole machine learning cycle, from data intake to continuous deployment, is integrated into MLOps pipeline and the efficiency of the pipelines will be evaluated. Total pipeline execution time and success rate for automation are the important metrics to be assessed.

# Monitoring

Continuous monitoring guarantees that the model operates at its best when produced. This evaluation will assess how well the system monitors performance. Metrics include the duration of identifying model problems and time taken to retrain in order to fix them.

# Conclusion

With an emphasis on automation, versioning, pipeline efficiency, and monitoring, the evaluation approach assesses the essential components of MLOps pipeline. By assessing these components, we can be assured that the system can scale according to increasing needs and preserve excellent performance.

# REFERENCES

Alla, S. and Adari, S.K. (2020). What Is MLOps? *Beginning MLOps with MLFlow*, [online] pp.79–124. doi:https://doi.org/10.1007/978-1-4842-6549-9\_3.

Deza, A. and Gift, N. (2021) *Practical Mlops*. sebastopol, california: O’Reilly Media, Inc.

Symeonidis, G., Nerantzis, E., Kazakis, A. and Papakostas, G.A. (2022). MLOps - Definitions, Tools and Challenges. *2022 IEEE 12th Annual Computing and Communication Workshop and Conference (CCWC)*. [online] doi:https://doi.org/10.1109/ccwc54503.2022.9720902.

Fessahaye, F., Perez, L., Zhan, T., Zhang, R., Fossier, C., Markarian, R., Chiu, C., Zhan, J., Gewali, L. and Oh, P. (2019). *T-RECSYS: A Novel Music Recommendation System Using Deep Learning*. [online] IEEE Xplore. doi:https://doi.org/10.1109/ICCE.2019.8662028.

Chen, H.-C. and Chen, A.L.P. (2005). A Music Recommendation System Based on Music and User Grouping. *Journal of Intelligent Information Systems*, [online] 24(2-3), pp.113–132. doi:https://doi.org/10.1007/s10844-005-0319-3.

Jordan, M.I. and Mitchell, T.M. (2020). Machine learning: Trends, perspectives, and prospects. *Science*, [online] 349(6245), pp.255–260. doi:https://doi.org/10.1126/science.aaa8415 (Accessed: 23 October 2024).

Treveil, M. *et al.* (2021) *Introducing Mlops: How to scale machine learning in the Enterprise*. Sebastopol, CA, california: O’Reilly Media, Inc.

Ebert, C., Gallardo, G., Hernantes, J. and Serrano, N. (2016). DevOps. *IEEE Software*, [online] 33(3), pp.94–100. doi:https://doi.org/10.1109/ms.2016.68 (Accessed: 24 October 2024).

John, M.M., Olsson, H.H. and Bosch, J. (2021). *Towards MLOps: A Framework and Maturity Model*. [online] IEEE Xplore. doi:https://doi.org/10.1109/SEAA53835.2021.00050.

Singla, A. (2023). Machine Learning Operations (MLOps): Challenges and Strategies. *Journal of Knowledge Learning and Science Technology ISSN: 2959-6386 (online)*, [online] 2(3), pp.333–340. doi:https://doi.org/10.60087/jklst.vol2.n3.p340 Accessed 28 Oct. 2024.

Liu, Z., Xu, W., Zhang, W. and Jiang, Q. (2023). An emotion-based personalized music recommendation framework for emotion improvement. *Information Processing & Management*, [online] 60(3), p.103256. doi:https://doi.org/10.1016/j.ipm.2022.103256 Accessed 28 Oct. 2024.

Wang, D., Zhang, X., Yin, Y., Yu, D., Xu, G. and Deng, S. (2023). Multi-View Enhanced Graph Attention Network for Session-Based Music Recommendation. *ACM Transactions on Information Systems*, [online] 42(1), pp.1–30. doi:https://doi.org/10.1145/3592853 (Accessed 28 Oct. 2024).

Kreuzberger, D., Kühl, N. and Hirschl, S. (2023). Machine Learning Operations (MLOps): Overview, Definition, and Architecture. *IEEE Access*, [online] 11, pp.1–1. doi:https://doi.org/10.1109/access.2023.3262138.

Granlund, T., Kopponen, A., Stirbu, V., Myllyaho, L. and Mikkonen, T. (2021). *MLOps Challenges in Multi-Organization Setup: Experiences from Two Real-World Cases*. [online] IEEE Xplore. doi:https://doi.org/10.1109/WAIN52551.2021.00019.

Hymel, S., Banbury, C., Situnayake, D., Elium, A., Ward, C., Kelcey, M., Baaijens, M., Majchrzycki, M., Plunkett, J., Tischler, D., Grande, A., Moreau, L., Maslov, D., Beavis, A., Jongboom, J. and Reddi, V.J. (2022). Edge Impulse: An MLOps Platform for Tiny Machine Learning. *arXiv:2212.03332 [cs]*. [online] Available at: https://arxiv.org/abs/2212.03332 [Accessed 4 Nov. 2024].

Hewage, N. and Meedeniya, D. (2022). MACHINE LEARNING OPERATIONS: A SURVEY ON MLOPS TOOL SUPPORT \*. *ArXiv*. [online] doi:https://doi.org/10.48550/arXiv.2202.10169.

Pradhan, A., Rathod, V., Jamdar, S. and Dangche, K. (2024). *Web Protection by Bitdefender*. [online] Ijeast.com. Available at: https://www.ijeast.com/papers/68-70 [Accessed 4 Nov. 2024].

Mishra, N., Gupta, R. and Raj, A. (2024). Music Recommendation System by Analyzing Facial Emotions Using Deep Neural Network. *SSRN Electronic Journal*. [online] doi:https://doi.org/10.2139/ssrn.4834085.

Chauhan, R., Ghanshala, K.K. and Joshi, R.C. (2018). *Convolutional Neural Network (CNN) for Image Detection and Recognition*. [online] IEEE Xplore. doi:https://doi.org/10.1109/ICSCCC.2018.8703316.

Susila, P. (2022). *Effective Use Of Mlops In Music Recommendation System MSc In Cloud Computing*. [online] *NCI*. Available at: https://norma.ncirl.ie/6476/1/pradeepmanoharansusila.pdf [Accessed 4 Nov. 2024].

Ruf, P., Madan, M., Reich, C. and Ould-Abdeslam, D. (2021). Demystifying MLOps and Presenting a Recipe for the Selection of Open-Source Tools. *Applied Sciences*, [online] 11(19), p.8861. doi:https://doi.org/10.3390/app11198861.

Ghantous, G. and Gill, A. (2017). *Association for Information Systems AIS Electronic Library (AISeL) DevOps: Concepts, Practices, Tools, Benefits and Challenges*. [online] PACIS. Available at: https://opus.lib.uts.edu.au/bitstream/10453/130066/1/DevOps-%20Concepts%20Practices%20Tools%20Benefits%20and%20Challenges.pdf [Accessed 4 Nov. 2024].

Banerjee, A., Chen, C.-C., Hung, C.-C., Huang, X., Wang, Y. and Chevesaran, R. (2020). *Challenges and Experiences with MLOps for Performance Diagnostics in Hybrid-Cloud Enterprise Software Deployments Challenges and Experiences with MLOps for Performance Diagnostics in Hybrid-Cloud Enterprise Software Deployments*. [online] USENIX. Available at: https://www.usenix.org/system/files/opml20\_paper\_banerjee.pdf [Accessed 4 Nov. 2024].

Shakti Sri, S., Sathya, S. and Pandiarajan, T. (2023). An Improved Music Recommendation System for Facial Recognition and Mood Detection. *ITM Web of Conferences*, [online] 56, p.01004. doi:https://doi.org/10.1051/itmconf/20235601004.

Subramanya, R., Sierla, S. and Vyatkin, V. (2022). From DevOps to MLOps: Overview and Application to Electricity Market Forecasting. *Applied Sciences*, [online] 12(19), p.9851. doi:https://doi.org/10.3390/app12199851.

Testi, M., Ballabio, M., Frontoni, E., Iannello, G., Moccia, S., Soda, P. and Vessio, G. (2022). MLOps: a Taxonomy and a Methodology. *IEEE Access*, [online] 10, pp.63606–63618. doi:https://doi.org/10.1109/access.2022.3181730.

# APPENDIX: PLES

# 

# APPENDIX: PROJECT MANAGEMENT

A screenshot of a computer

Description automatically generated