

New

Ministry

Position Name (30 characters)

Design: Identify Job Duties and Value**Job Purpose and Organizational Context**

Why the job exists:

The Machine Learning Operations (MLOps) Specialist is an integral part of the Technology and Innovation Artificial Intelligence (AI) Program team, providing technical support for the deployment, maintenance, and optimization of AI solutions. The MLOps Specialist brings expertise in data engineering, DevOps practices, and Generative AI technologies to the Government of Alberta (GoA). This role ensures the high reliability and performance of AI systems in production environments, aligning with the GoA's AI Strategy and aiding in the implementation of the GoA AI Roadmap. The MLOps Specialist advises on best practices and standards for deploying, monitoring, and scaling AI solutions. Additionally, the MLOps Specialist collaborates closely with third-party AI solution providers contracted by the GoA, ensuring compliance with GoA's MLOps guidelines and requirements. Responsibilities include designing and implementing data pipelines, workflows, and infrastructure for both traditional and generative machine learning models. Projects involve utilizing cloud services, automation tools, and advanced programming skills to facilitate continuous integration and delivery of AI solutions. The MLOps Specialist also supports Data Scientists and Software Engineers in developing, testing, and refining machine learning models and collaborates with unit staff and business units across the GoA to optimize AI solutions and resolve issues.

Responsibilities

Job outcomes (4-6 core results), and for each outcome, 4-6 corresponding activities:

technologies such as Docker, orchestration tools like Kubernetes, and Azure App Services.

2. Infrastructure Management: Collaborate with DevOps team to manage and optimize infrastructure resources for machine learning workloads, ensuring scalability, reliability, and cost-efficiency.
3. Pipeline Automation: Develop and maintain automated pipelines for model training, evaluation, and deployment, integrating with version control systems and CI/CD pipelines.
4. Monitoring and Alerting: Implement advanced monitoring and alerting systems to enhance observability of deployed models in production. Track model performance, detect data drift, and monitor system health using best practices and tools from Microsoft Azure, Google, Amazon, etc.. This proactive approach ensures timely response to issues and maintains compliance with organizational standards.
5. Model Versioning and Governance: Establish version control and governance processes for machine learning models, ensuring traceability, reproducibility, and compliance with regulatory requirements.
6. Security and Compliance: Implement security best practices for handling sensitive data and ensure compliance with data protection regulations like FOIP, GDPR and HIPAA.
7. Collaboration and Documentation: Work closely with cross-functional teams, including data scientists, software engineers, and product managers, to understand requirements, provide technical guidance, and document processes and best practices.

Problem Solving

Typical problems solved:

A Machine Learning Operations (MLOps) Specialist focuses on the operational aspects of deploying, managing, and scaling machine learning models in production environments. Some typical problems they solve include:

- Model Deployment: Ensuring that machine learning models can be deployed reliably and efficiently into production environments. This includes automating the deployment process and integrating with existing IT infrastructure.
- Model Monitoring: Implementing systems to continuously monitor model performance in production. This involves tracking metrics such as accuracy, latency, and throughput to ensure models are performing as expected.
- Model Versioning: Managing different versions of models to ensure that updates or rollbacks can be performed smoothly. This involves tracking changes, maintaining historical versions, and ensuring compatibility with existing systems.
- Scaling: Ensuring that models can scale to handle varying loads and data volumes. This might involve optimizing resource allocation, load balancing, and distributed computing strategies.
- Data Management: Handling data pipelines and workflows to ensure that models have access to the necessary data for training and inference. This includes managing data ingestion, preprocessing, and storage.
- Reproducibility: Ensuring that experiments and model training can be reproduced accurately. This involves managing code, data, and environment dependencies to guarantee that results are consistent.
- Automation: Automating various aspects of the ML lifecycle, including data preprocessing, model training, testing, and deployment. This helps to streamline operations and reduce manual intervention.
- Security and Compliance: Implementing security measures to protect sensitive data and ensure compliance with relevant regulations and standards. This includes managing access controls, auditing, and data encryption.
- Incident Management: Responding to and resolving issues that arise in the production environment, such as model performance degradation, system outages, or data inconsistencies.
- Collaboration and Communication: Facilitating communication between data scientists, machine learning developers and other stakeholders to ensure that ML projects are aligned with business goals and that everyone is informed about

progress and issues.

- Cost Management: Managing the costs associated with running and maintaining machine learning models, including optimizing resource usage and exploring cost-effective solutions.

By addressing these problems, MLOps Specialists help ensure that machine learning initiatives are successful, scalable, and sustainable in real-world applications.

Types of guidance available for problem solving:

Best Practices and Frameworks: Following established frameworks like Microsoft's Azure ML Ops, Google's MLOps, etc. or the MLflow framework can provide structured approaches to managing ML workflows.

Industry Best Practices: Adhering to best practices for model deployment, monitoring, and scaling, such as those outlined in the MLOps community or cloud service provider MLOps documentation.

Documentation and Guidelines: Detailed guides and tutorials provided by cloud service providers like Azure, AWS, or Google Cloud Platform (GCP) on their MLOps tools and services.

Open-Source Tools: Documentation and user guides for open-source MLOps tools like Kubeflow, MLflow, or Metaflow.

Educational Resources: Online Courses and Tutorials.

Webinars and Workshops: Industry webinars, workshops, and meetups can provide practical insights and hands-on experience.

Community and Forums: Engaging with communities such as the MLOps community, specialized LinkedIn groups, etc. Participating in forums like Stack Overflow or GitHub discussions to get help and advice from peers and experts.

Industry Experts: Leveraging advice from experienced professionals through networking or speaking engagements.

Tools and Platforms: Utilizing tools for automated model deployment and monitoring. Implementing monitoring solutions for real-time insights into model performance and system health.

Research and Case Studies: Reviewing academic research on MLOps challenges and solutions. Analyzing case studies from other organizations to understand how they tackled similar problems.

Internal Documentation: Documenting and sharing solutions and best practices within the organization to build a knowledge base.

Experimentation and Testing: Running experiments to compare different approaches and determine the best solution. Developing prototypes to test new ideas and approaches before full-scale implementation.

Combining these types of guidance can help MLOps Specialists effectively address challenges and continuously improve their processes and systems.

Direct or indirect impacts of decisions:

Direct Impacts:

1. Improved Model Performance

- **Deployment:** Efficient deployment processes ensure that models are delivered into production with minimal issues, directly improving their performance and reliability.
- **Monitoring and Feedback:** Continuous monitoring and feedback mechanisms can lead to prompt adjustments and refinements, directly enhancing model accuracy and performance.

2. Operational Efficiency

- Automation: Implementing automated pipelines for model training, deployment, and monitoring reduces manual effort and operational overhead, leading to more efficient use of resources.
- Resource Management: Optimizing resource usage (e.g. compute and storage) can lead to cost savings and better performance, directly impacting the organization's operational budget.

3. Enhanced Collaboration

- Cross-Functional Communication: Facilitating collaboration between Data Scientists, ML Developers and other stakeholders ensures that ML projects align with business objectives, improving project outcomes and efficiency.
- Knowledge Sharing: Sharing insights and best practices across teams can lead to better problem-solving and more effective ML solutions.

4. Risk Mitigation

- Incident Management: Rapidly addressing issues such as model drift or system failures minimizes disruptions and risks associated with production ML systems.
- Security: Implementing security measures helps protect sensitive data and ensures compliance, reducing the risk of data breaches and regulatory fines.

Indirect Impacts

1. Business Outcomes

- Decision-Making: By ensuring that ML models are reliable and perform as expected. MLOps Specialists indirectly support better decision-making based on accurate and timely data insights.
- Customer Experience: Improved ML systems can lead to enhanced customer experiences through personalized recommendations, better service quality, or more efficient processes.

2. Innovation and Growth

- Speed to Market: Efficient MLOps practices enable faster deployment of new models and features, fostering innovation and helping the organization stay competitive.
- Scalability: Effective scaling solutions support the growth of ML applications, allowing the organization to handle increased data and user demands.

3. Organizational Culture

- Collaboration Culture: Promoting a culture of collaboration and knowledge sharing can enhance overall team dynamics and foster a more cohesive work environment.
- Continuous Improvement: By embedding best practices and continuous improvement processes, MLOps Specialists contribute to a culture of excellence and ongoing learning.

4. Reputation and Trust

- Reliability: Consistently delivering reliable ML solutions builds trust with stakeholders and customers, enhancing the organization's reputation.
- Compliance: Adhering to regulatory standards and best practices in data handling and security reinforces the organization's commitment to ethical practices.

In summary, the role of an MLOps Specialist has both immediate and ripple effects on various aspects of an organization. Their interactions and the quality of their work can influence model performance, operational efficiency,

and broader business and cultural outcomes.

Key Relationships

Major stakeholders and purpose of interactions:

Works closely with the Director/Manager of AI, and ML Developers to ensure what is being deployed is being deployed and supported correctly. Works closely with Business Stakeholders to ensure the desired solution is being deployed as required.

Required Education, Experience and Technical Competencies

Education Level	Focus/Major	2nd Major/Minor if applicable	Designation
Bachelor's Degree (4 year)	Engineering	Science	

If other, specify:

Computer Science, Engineering or related

Job-specific experience, technical competencies, certification and/or training:

Education/Certifications

- Bachelor's degree or higher in Computer Science, Engineering, Mathematics, or related field.
- Certifications in cloud platforms (e.g. Microsoft Azure (preferred), GCP, etc.).

Experience:

- At least 6 years of proven experience deploying and managing machine learning models in production environments.
- At least 2 years of experience working with generative AI models and frameworks such as Langchain, huggingface and be familiar with large language models such as LLAMA, ChatGPT, and Claude etc.
- Experience in managing the entire product life cycle, from planning and development to deployment and updates.
- Knowledge of data management methods, principles and techniques.
- Strong analytical skills with the ability to collect, organize, analyze and disseminate significant amounts of information with attention to detail, accuracy and data quality.
- Strong problem-solver, with the ability to handle ambiguous and open-ended problems.
- Strong human relations and communication (written, verbal, and listening) skills will be necessary to effectively communicate across disciplines and stakeholders.

Technical Competencies:

- Proficiency in programming languages such as Python, and experience with libraries like TensorFlow, PyTorch, Scikit-learn, Flask/Quart, FastAPI, and Django.
- Strong understanding of containerization (e.g. Docker), orchestration (e.g. Kubernetes), and cloud platforms (e.g. Microsoft Azure (preferred), GCP, etc.).
- Experience with version control systems (e.g. Git), CI/CD pipelines, and infrastructure as code tools (e.g. Bicep (preferred), Terraform).
- Familiarity with monitoring and logging tools and proficiency in scripting languages (e.g. Bash, PowerShell).
- Knowledge of data engineering tools to conduct ETL, ELT, etc.
- Experience with MLOps platforms and tools such as MLflow, Kubeflow, etc.
- Knowledge of machine learning techniques and algorithms, with experience in model training and evaluation.
- Familiarity with distributed computing and parallel processing techniques.
- Knowledge of working with security and networking team at an enterprise level.
- Knowledge of managing data at scale using Spark and DataBricks.
- Knowledge of web app development with JavaScript and debugging tools using NodeJS.

Behavioral Competencies

Pick 4-5 representative behavioral competencies and their level.

Competency	Level					Level Definition	Examples of how this level best represents the job
	A	B	C	D	E		
Systems Thinking	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	Integrates broader context into planning: • Plans for how current situation is affected by	Looks for the best solution and prioritizes the execution that will have the least impact on

		<p>broader trends</p> <ul style="list-style-type: none"> • Integrates issues, political environment and risks when considering possible actions • Supports organization vision and goals through strategy • Addresses behaviours that challenge progress 	<p>stakeholders environments while keeping future opportunities for enhancements in mind.</p>
Creative Problem Solving	<input type="radio"/> <input type="radio"/> <input checked="" type="radio"/> <input type="radio"/> <input type="radio"/>	<p>Engages the community and resources at hand to address issues:</p> <ul style="list-style-type: none"> • Engages perspective to seek root causes • Finds ways to improve complex systems • Employs resources from other areas to solve problems • Engages others and encourages debate and idea generation to solve problems while addressing risks 	<p>Work with a variety of people including cloud, visualization, clients, external vendors and executives to properly execute the vision and priorities that have a direct impact on Albertans.</p>
Agility	<input type="radio"/> <input type="radio"/> <input checked="" type="radio"/> <input type="radio"/> <input type="radio"/>	<p>Identifies and manages required change and the associated risks:</p> <ul style="list-style-type: none"> • Identifies alternative approaches and supports others to do the same • Proactively explains impact of changes • Anticipates and mitigates emotions of others • Anticipates obstacles and stays focused on goals • Makes decisions and takes action in uncertain situations and creates a backup plan 	<p>Must be flexible and capable of managing change in constantly evolving environments as new policies and technologies. Must work collaboratively with various teams to ensure best possible support is provided.</p>
Drive for Results	<input type="radio"/> <input type="radio"/> <input checked="" type="radio"/> <input type="radio"/> <input type="radio"/>	<p>Takes and delegates responsibility for outcomes:</p> <ul style="list-style-type: none"> • Uses variety of resources to monitor own performance standards • Acknowledges even indirect responsibility • Commits to what is good for Albertans even if not immediately accepted • Reaches goals 	<p>Focused in achieving the near term and long term goals of the group making sure the team is motivated and ready to take new challenges as they come along.</p>

		consistent with APS direction	
--	--	----------------------------------	--

Benchmarks

List 1-2 potential comparable Government of Alberta:

[Benchmark](#)

Assign

The signatures below indicate that all parties have read and agree that the job description accurately reflects the work assigned and required in the organization.

Employee Name

Date yyyy-mm-dd

Employee Signature

Supervisor / Manager Name

Date yyyy-mm-dd

Supervisor / Manager Signature

Director / Executive Director Name

Date yyyy-mm-dd

Director / Executive Director Signature