

RAKEZ Case Study: Deploying and Monitoring a Lead Scoring Model

Role: Machine Learning Engineer

Duration: 5–7 days to submit

Format: PDF or PPT (with optional code snippets or visuals)

Background

RAKEZ has developed a lead scoring model to predict the probability of a new lead converting into a paying customer. The model has shown promising results in offline validation and is now ready to be deployed in a production environment.

As an ML Engineer at RAKEZ, you will play a crucial role in deploying, monitoring, and maintaining this model in a live setting. Your ability to ensure model performance, reliability, and transparency in a real-time environment will directly influence business outcomes.

Case Objective

You are required to design a robust solution that enables:

1. Production deployment of the lead scoring model
2. Online testing strategy (e.g., A/B testing, shadow deployment)
3. Ongoing monitoring and alerting of model performance
4. Reproducibility, automation, and retraining readiness

Your Deliverables

Please submit a short case study proposal (5–10 slides or pages) that addresses the following:

1. Deployment Strategy

- How would you deploy this model in a production environment (assume Python and Databricks are available)?
- What frameworks or tools (e.g., MLflow, Docker, REST API) would you use?
- How would you handle model versioning, auditability, and rollback?

2. Online Testing Approach

- Propose a method for validating the model in production (e.g., A/B testing or shadow deployment).
- What metrics would you track to determine success?
- How would you ensure that testing doesn't negatively impact business operations?

3. Monitoring Plan

- What metrics would you track for:
 - Data drift
 - Prediction drift
 - Latency and throughput
 - Business performance (e.g., conversion rates)
- How would you implement alerting and logging?
- The sales team reports that the model's lead scores are no longer helping prioritize leads effectively. What steps would you take to investigate whether the issue lies in the model, the data, or the integration?

4. Automation, Reproducibility & Retraining

- How would you ensure reproducibility (e.g., consistent data snapshots, version control)?
- What CI/CD workflows would you build for safe deployment?
- When and how would you trigger model retraining?

Optional (Bonus Points)

- Include a lightweight monitoring dashboard sketch or code snippet
- Suggest how you'd expose model predictions to the CRM or business dashboards
- Outline a feedback loop using user or sales team inputs

Submission Guidelines

- Format: PDF, PPT, or Markdown document
- Diagrams are encouraged (architecture, pipelines, dashboards)
- Optionally include a GitHub link or Jupyter Notebook if applicable