DS/CS553 – Machine Learning Development and Operations (MLOps)

Case Study 2

The purpose of this case study is to deepen your understanding of deployment and automated recovery of machine learning applications. You will expand upon the task from Case Study 1 by deploying your products to a virtual machine environment and implementing automated recovery mechanisms. Professor Paffenroth has provided virtual machines (VMs) for this assignment. You will begin by establishing SSH access to these VMs, setting up the necessary environments, and deploying both products you developed in Case Study 1. Finally, you will configure your deployments to automatically recover and restart in case of failures.

Like with Case Study 1, this assignment will be done in groups of 1-3 people each. The total number of points for the case study will be 100 (with an additional 10 points of extra credit), distributed as shown below.

Deliverables

1. Virtual Machine Setup (**20 points**)
    a. SSH Access Setup (*10 points*)
    Establish SSH access to the virtual machine provided for your group by Professor Paffenroth. You can refer to the set-up documentation provided in the notes from classes 7 and 8. Make sure to replace the public key authorized on the machine with your own, so that only your group has access to it! Write up a brief outline of the steps involved in setting up the SSH access, including any required configurations and credentials.
    b. Environment Configuration (*10 points*)
    Set up the virtual environment and all necessary dependencies to run both the API-based and local products developed in Case Study 1. Document the setup process, including the installation of all required libraries, environment variables, and configurations needed to run the applications.
2. Deployment of Products on Virtual Machines (**30 points**)
    a. Deployment of API-based Product (*15 points*)
    Deploy the API-based product from Case Study 1 on the virtual machine. Ensure that the product is fully functional and accessible via the provided virtual machine. Document any modifications made to the deployment process to suit the VM environment.
    b. Deployment of Locally Executed Product (*15 points*)
    Deploy the locally executed product from Case Study 1 on the virtual machine. Ensure that the product is fully operational and behaves as expected in the VM environment.

3. Automated Deployment (**30 points**)
    a. Script for Automated Deployment (*15 points*)
    Implement a script that does an automatic deployment of your product. This would involve copying personal public keys, adding them to the list of authorized keys on the virtual machine, and removing the standard student-admin default key. The script should then install the required virtual environment, clone the GitHub repository where your products reside, and deploy them to the virtual machine. Provide the code and documentation for your script, explaining each step in detail.
    b. Resilience Testing (*15 points*)
    Sometimes, the virtual machines you are working on can fail and be completely wiped. *When* this happens, verify that the recovery mechanism successfully restores the products to a functional state. Document the outcome and any challenges encountered.
4. Project Report—2-3 pages long (**20 points**)
    a. Group Members' Names (*1 point*)
    b. Virtual Machine Setup Process (*5 points*)
    Outline the steps taken to set up SSH access to the virtual machines and the virtual environment used.
    c. Deployment Process (*5 points*)
    Explain how you deployed both products on the virtual machines, including any challenges and how they were addressed.
    d. Automated Recovery (*5 points*)
    Document your experience with the automated recovery process, including its effectiveness, potential improvements, and any limitations observed during testing.
    e. Additional Insights, Challenges, and Future Improvements (*4 points*)
    Reflect on this case study, discussing any additional challenges faced, lessons learned, and potential future improvements to the deployment, monitoring, or recovery processes.
5. Extra Credit (+**10 points**)
    a. Automatic Server Recovery (*5 points*)
    Develop a system that periodically checks and detects when the virtual machine goes down and triggers this recovery process automatically (i.e., without having to manually access the VM).
    b. Red-Teaming (*5 points)*
    Remember that you should replace the public key authorized on the machine with your own, so that only your group has access to it. If at any point after 12:00 AM on September 30th, 2025 (halfway through this assignment), you are able to access another group's virtual machine (because they did not replace the default key with their own) and provide a screenshot of doing so with the date and time, you will receive 5 points of extra credit, **provided your machine has not been accessed by another team**. *Please respect both your teammates (by letting them know that you have accessed their machine and make no changes to their*

machine!) and the assignment (do not just fake your screenshot). ***Doing anything other than accessing the virtual machine and taking a screenshot (including, but not limited to, changing or removing files, making configuration changes, leaving a backdoor, any attempt to access the underlying hardware, etc.) will result in no extra credit being awarded. Also, be aware that port scanning will almost certainly get your IP banned by the WPI security system.***

All materials should be submitted on canvas.wpi.edu.

Links to materials that will help with the case study:

- https://github.com/rcpaffenroth/DSCS553_example