

الجمهورية الجزائرية الديمقراطية الشعبية
People's Democratic Republic of Algeria
وزارة التعليم العالي و البحث العلمي
Ministry of Higher Education and Scientific Research
المدرسة العليا للإعلام الآلي 8 ماي 1945 - سيدي بلعباس
Higher School of Computer Science
8 Mai 1945 - Sidi Bel Abbès



Final Year Graduation Project Thesis

To obtain the diploma of **Engineer's Degree**

Field of Study: **Computer Science**

Specialization: **Artificial intelligence and Data science**

Theme

Dependency-Aware Microservice Autoscaling with Graph Neural Networks and Reinforcement Learning

Presented by

Slimane MEHARZI

Mohamed Amine FELLAH

Defended on: **October, 2025**

In front of the jury composed of

Mr. [Jury Member Name]

Mr. Abdelhamid MALKI

Mr. [Jury Member Name]

President of the Jury

Thesis Supervisor

Examiner

Academic Year: 2024/2025

Abstract

Traditional autoscaling for microservice architectures is often inefficient as it is reactive and ignores the complex, graph-like dependencies between services, leading to costly overprovisioning or performance degradation. This thesis addresses these limitations by presenting a novel, dependency-aware autoscaling framework for Kubernetes that integrates a Graph Neural Network (GNN) with a Reinforcement Learning (RL) agent. The GNN processes metrics from individual services and the traffic between them to generate a holistic, real-time embedding of the system’s state, which the RL agent then uses to learn a sophisticated, system-wide scaling policy. Implemented and validated on a Kubernetes cluster using the Online Boutique application with an Istio and Prometheus monitoring stack, the agent successfully learned a complex control policy. A critical finding reveals the extreme sensitivity of the learning process to the mathematical formulation of the reward signal, a key challenge in reward engineering for live systems. The primary contribution is the complete design, implementation, and critical analysis of this end-to-end GNN-RL framework, offering valuable lessons for the practical application of reinforcement learning to dynamic resource management.

Keywords— Microservices, Autoscaling, Cloud Computing, Resource Management, Reinforcement Learning.